Worker Mobility and the Diffusion of Radical Technologies

Stephan Hobler*
London School of Economics

13th November 2025

Please click here for the latest version.

Abstract

The spread of new, transformative technologies often relies on specialized knowledge among workers and managers. When the required expertise is scarce, human capital and worker mobility can become bottlenecks to technology diffusion. To study these dynamics, I develop a theory in which firms and workers accumulate technology-specific expertise through mutual learning, and worker mobility is subject to search frictions. I calibrate the model to the diffusion of predictive AI among U.S. firms, matching empirical patterns of technology adoption and worker flows using comprehensive microdata from LinkedIn. Worker mobility emerges as a key driver of diffusion. When a new technology is introduced, adoption is initially slow due to the lack of experienced workers and concentrated only among the most productive firm-worker pairs. Diffusion then accelerates through the poaching of workers from early adopters, generating an S-shaped diffusion curve. Aggregate output follows a J-curve, with productivity gains taking time to materialize. A counterfactual matched to European-style labor markets with low mobility and long job tenures reveals a trade-off: the European economy delivers modestly higher steady-state output, but slows adoption by two-thirds and reduces welfare gains from the new technology by one-third—potentially contributing to transatlantic productivity gaps in technology-intensive sectors.

Keywords: technology adoption; radical technology; vintage human capital; worker mobility

JEL Codes: O33; E22; D83; J24

*Department of Economics. Contact: s.j.hobler@lse.ac.uk. Website: https://sjhobler.github.io. I am indebted to Matthias Doepke, Wouter Den Haan, and Joe Hazell for invaluable guidance and support. I also thank Marco Bellifemine, Michael Cai, Francesco Caselli, Adrien Couturier, Maarten De Ridder, Johannes Matt, Benjamin Moll, Aditya Soenarjo, and Silvana Tenreyro for valuable comments and discussions.

1 Introduction

Economic growth often unfolds in waves, each triggered by a radical technological breakthrough followed by continuous incremental improvements (Bresnahan & Trajtenberg, 1995; Mokyr, 2002; Schumpeter, 1939). A defining feature of many such technologies is that their adoption and implementation depend critically on specialized knowledge. When expertise in these technologies is scarce, it can become a bottleneck to their diffusion. What drives the spread of technologies that are increasingly dependent on scarce human capital? Understanding the underlying mechanisms is crucial to determine whether and how policy can facilitate a faster transition and maximize potential welfare gains from the new technology.

This paper stresses the central role of labor markets and specifically worker mobility in shaping the diffusion of radical technologies—that is, technologies that offer substantial productivity gains, but require the accumulation of specialized knowledge for successful implementation. Because firms and workers gradually accumulate this knowledge through experience, worker mobility becomes a critical channel for its transfer, particularly during early adoption when implementation uncertainty is highest. Beyond technical expertise, they bring tacit knowledge about adapting existing business practices to new technologies and help firms avoid costly mistakes by sharing lessons from failed experiments at previous employers.

The core objective of this paper is to formalize this worker mobility channel and characterize its importance for the diffusion of technologies. To this end, I make three main contributions. First, I develop a general framework of technology adoption and knowledge diffusion through worker flows, in which firms and workers jointly accumulate technology-specific expertise, and existing knowledge is only partially transferable when upgrading to newer vintages. Second, I apply the framework to a particular technology: Artificial Intelligence (AI) adoption in U.S. labor markets. I quantify the role of worker reallocation and poaching in its diffusion, matching joint patterns of technology adoption and worker flows using comprehensive resume data to discipline the model. Third, I show how labor market institutions can have sizable effects on adoption patterns and aggregate welfare in a counterfactual exercise re-calibrated to low mobility and long tenure jobs as seen in many EU labor markets.

The paper begins by documenting the correlation of worker flows with the adoption of radical new technologies. To establish descriptive patterns, I focus on a particular application of such a technology:

the adoption of predictive AI across firms in U.S. metropolitan areas. Using comprehensive resume data covering the near-universe of LinkedIn profiles, I classify firms as adopters and non-adopters based on AI-related keywords in workers' job titles and job descriptions. I then examine whether newly adopting firms disproportionately hire workers from previous adopters. Two key patterns emerge from the data. First, I find that the excess probability of hiring a worker from an AI-adopter relative to a non-adopter jumps by 15% at the time of adoption—revealing the demand for specialized external human capital when implementing a new technology. Second, hiring concentrates among experienced rather than junior workers initially, suggesting that worker mobility serves as a channel for transferring practical implementation knowledge and not exclusively technical skills.

Motivated by these empirical patterns, I develop a theoretical framework to study how worker mobility and learning jointly shape the diffusion of new, path-breaking technologies. The economy is endowed with a set of technology vintages, where the new generation offers higher productivity, but requires a higher per-period fixed cost. There are two types of agents—workers and entrepreneurs—who meet in a frictional labor market. Workers are heterogeneous in both their level of human capital, which they accumulate specific to a technology vintage, and their employment status. Entrepreneurs differ in their productivity, which is specific to the technology vintage they currently operate. The setup distinguishes between two forms of technological progress: gradual productivity improvements within a given vintage versus discrete jumps to entirely new technology vintages. Entrepreneurs are either matched with a single worker or operate solo and post vacancies in search of a profitable match. In both states, entrepreneurs produce output, but employing a worker offers two advantages: it enhances production through complementarities in the worker's human capital and the entrepreneur's productivity, and it enables mutual learning through the gradual transfer of knowledge.[1]

An entrepreneur-worker match jointly decides whether to upgrade to a new technology vintage. When upgrading, existing knowledge is only partially transferable from the older to the newer vintage. Combined with the mutual learning process, this partial transferability creates a central role for experienced workers and their reallocation in driving the diffusion of a new technology. Experienced workers

[1]This creates a trade-off between learning and the opportunity cost of foregone production. Production complementarities favor matching high productivity entrepreneurs with high human capital workers while learning gains favor matching agents with different knowledge levels.

contribute through two channels. First, their accumulated technology-specific experience directly enhances production. Second, they facilitate the transfer of knowledge, allowing entrepreneurs to more quickly acquire technology-specific knowledge through mutual learning and thereby incentivizing adoption through external hiring.

I study the transition path of the economy when a new technology vintage is introduced. When calibrated to U.S. labor market data and using predictive AI as an application, the quantitative model delivers a characteristic S-shaped diffusion curve—taking approximately 3 years to reach 5% penetration and 12 years to reach 50%. Early in the transition, the lack of experienced workers constrains adoption. Adoption is therefore more costly initially as it occurs internally with the firm's existing inexperienced workers, and only the most productive matches adopt. Subsequently, as more entrepreneurs match with experienced workers, the human capital bottleneck relaxes and diffusion accelerates. The process eventually decelerates as agents in the economy become more knowledgeable in the new technology and the labor market sorts high-human-capital workers into high-productivity firms.[2] Then, successful poaching becomes more difficult and costly. Aggregate output follows a J-curve, with output lagging behind adoption as firms and workers incrementally build up their productivity and human capital in the new technology vintage, consistent with prior empirical and theoretical literature on general purpose technologies (Brynjolfsson et al., 2021; David, 1990; Helpman & Trajtenberg, 1994).

Central to these dynamics, access to workers with experience in the new vintage is a binding constraint in the calibrated version of the model: while nearly all firms would adopt if they had access to a worker experienced in the new technology, only 50% of potential adoption-enabling matches are actually formed early in the transition—consistent with survey evidence on skill constraints (European Commission, 2020; Rammer & Schubert, 2022). To understand how this constraint relaxes over time, I decompose adoption into its distinct channels. Consistent with the data, internal adoption without external hiring constitutes the smallest fraction. Only approximately 5% of all adoption occurs in the initial phase without external hiring. The remaining 95% occurs through external hiring, but the composition shifts over the transition: employment-to-employment (E-E) flows dominate early when the need for knowledge trans-

[2]This shares the insight with Engbom (2023) that new technologies are more valuable when individuals had less time to become well-matched and workers are cheaper to poach. Different from that paper, these dynamics correspond to the beginning of the transition—a period of slow growth—while the joint process of sorting and learning drives the subsequent high growth.

fer is most acute, while slower unemployment-based reallocation becomes prevalent later as the pool of experienced workers grows.[3]

Thus, the speed of the transition crucially hinges on the strength of the poaching channel: Without E-E flows, diffusion relies on the slower processes of unemployment reallocation and internal adoption, which lacks knowledge spillover benefits. I discipline the strength of poaching using the excess hiring probabilities estimated using the resume data.

The prominence of worker mobility in the calibrated model suggests that labor market institutions are important for diffusing radical technologies. To examine this further, I apply the framework to European-style labor markets and contrast results with the U.S. baseline. This comparison is particularly relevant given the diverging economic performance between the U.S. and Europe. Draghi (2024) finds that this divergence can be primarily attributed to the Information and Communication Technologies (ICT) sector and the use of modern technologies—precisely where my model predicts labor mobility constraints would bind most severely.[4] Specifically, I re-calibrate labor market parameters to match lower mobility rates and longer job duration as seen in many EU labor markets (Borowczyk-Martins, 2025), thereby capturing cross-country differences in labor market institutions in reduced form. While aggregate output in the old steady-state is modestly higher relative to the U.S. calibration, the new technology's diffusion slows down substantially. The time taken to reach 50% penetration of the new technology increases by two-thirds in the less dynamic (EU) labor market, and welfare gains from the introduction of the radical new technology are lower by one-third. This reveals a fundamental trade-off between labor market stability in the steady-state and technological dynamism during the transition. Moreover, these results highlight that labor market institutions can have sizable effects on adoption speeds and welfare, proposing a quantitative mechanism for the transatlantic productivity gaps.[5]

In the final part of the paper, I examine whether policy can and should accelerate diffusion, taking labor market frictions as given. I study a specific policy instrument: a time-restricted subsidy paid to adopters of the new technology. I show that while temporary subsidies for early adopters increase welfare,

[3] Unemployment in the model represents various slow-moving reallocation factors including quits for personal reasons or amenity values. These transitions from adopted firms create a gradual increase in the supply of experienced workers.

[4] In traditional sectors such as manufacturing EU economies are on par or even outperform the U.S. (Draghi, 2024).

[5] This complements general arguments about dynamism (Schoefer, 2025) by identifying the precise channel through which mobility matters: the transfer of implementation knowledge from those who adopted previously.

persistent subsidies can reduce it. This finding emerges from two offsetting inefficiencies. First, matches do not internalize the positive spillover effects their adoption creates through the mobility of their worker. This externality is strongest early on when most entrepreneurs have not yet adopted. Second, search frictions induce congestion effects in the labor market, which peak later in the transition when competition for workers intensifies. An expiring subsidy balances these externalities by alleviating the initial under-adoption problem while expiring before congestion effects intensify.

Literature. This paper adds to three strands of literature. First, I build on the knowledge diffusion literature (Benhabib et al., 2021; Buera & Oberfield, 2020; Lucas & Moll, 2014; Luttmer, 2007; Nelson & Phelps, 1966; Perla & Tonetti, 2014; Perla et al., 2021; Sampson, 2015). While these foundational papers study how productivity-enhancing knowledge propagates through the economy and shapes long-run growth patters, they are largely agnostic about the specific channel of diffusion. Recent work by Jingnan (2023) and Koike-Mori et al. (2024) examine knowledge diffusion through the reallocation of *inventors* around a balanced growth path.[6] In contrast, I stress the importance of worker mobility of *implementors* in the adoption of a radical new technology, i.e. new technologies which require specialized knowledge to be accumulated to effectively implement them. Moreover, I characterize the full transition dynamics over the lifecycle of a technology and show how labor market institutions fundamentally shape these diffusion patterns in the presence of search frictions. This transition focus generates both the S-shaped adoption curves and J-curve productivity patterns, while revealing how labor market institutions mediate the trade-off between stability in the steady-state and technological dynamism during the transition.

Second, the paper extends the literature on coworker learning (Adenbaum et al., 2024; Herkenhoff et al., 2024; Jarosch et al., 2021).[7] While these papers demonstrate how within-team knowledge spillovers shape individual productivity, wage dynamics, and labor market inequality, they abstract from firm-level technology adoption decisions and do not differentiate between incremental and radical innovations—two central features in my analysis. Without the radical technology component and firm adoption decisions, my model reduces to a framework similar to Herkenhoff et al. (2024).

[6]A different strand of the literature links worker mobility and in particular job creation and job destruction to firm dynamics and the overall business dynamism (Davis & Haltiwanger, 2014; Decker et al., 2016; Hopenhayn & Rogerson, 1993).

[7]The importance of teams in research is shown in Akcigit et al. (2018), Jaravel et al. (2018) and Jones (2009).

Third, the paper contributes to work on vintage human capital (Adão et al., 2024; Chari & Hopenhayn, 1991; Jovanovic & Nyarko, 1996; Violante, 2002), General Purpose Technologies (GPT) (Aghion et al., 2002; Helpman & Trajtenberg, 1994, 1996; Jovanovic & Rousseau, 2005), and technological revolutions (Atkeson & Kehoe, 2007; Caselli, 1999). While the vintage human capital literature largely focuses on generational changes in the supply of skilled labor in an Overlapping Generations framework, I emphasize worker mobility as an important channel for accessing skilled labor and model the interaction between worker and firm-side vintage knowledge.[8] My concept of radical technology draws parallels to General Purpose Technologies while its introduction can be interpreted as a technological revolution. The focus in the present paper is on the *diffusion* of GPTs which rely on specialized knowledge.[9]

Outline. Section 2 documents stylized facts that motivate the theory. Section 3 develops the theory. Its predictions are quantified in Section 4 and applied to the case of the growing EU-U.S. divide. Section 5 studies a particular subsidy on adoption and derives optimal policy. Section 6 concludes.

2 Empirical motivation: Worker flows and AI adoption

Before turning to the theoretical framework for a general class of technologies, I examine descriptive patterns for the emergence of a particular example of a radical technology: Artificial Intelligence (AI). Specifically, the context is predictive AI, which was the predominant form of AI adopted during the 2010s.

Survey evidence across multiple countries establishes that access to skilled labor acts as a key constraint on AI adoption: approximately 50% of firms in the European enterprise survey on the use of technologies cite skill shortages as a major obstacle to AI adoption, comparable to concerns about financial costs and investment risks (European Commission, 2020). The Mannheim Innovation Survey (Rammer & Schubert, 2022) similarly finds that 40-60% of large knowledge-intensive firms use hiring to access third-party know-how. These findings underscore how access to skilled labor shapes technology adoption decisions, which will be a central element in the theoretical framework.

[8] Evidence of skill obsolescence and vintage human capital is shown in Hombert and Matray (2024) and Kogan et al. (2021).

[9] The focus on the *diffusion* is complementary to Acemoglu et al. (2022a), who study where radical (creative) innovations come from—stressing the importance of the human capital of newer generations of managers.

Building on this evidence, I explore the specific mechanism through which firms acquire needed expertise: worker poaching from prior adopters. I document the correlation between worker flows and AI adoption patterns in U.S. firms in digitally-intensive sectors.[10] To analyze whether firms' hiring practices change around the time of adoption and, if so, for which workers and from which firms, I adapt the concept of *excess probabilities* from Cestone et al. (2023). This exercise provides motivating evidence for the model's mechanism and a key empirical moment for calibration.

2.1 Data

The primary dataset is Revelio Labs, a workforce intelligence company that provides harmonized resume data covering the near-universe of LinkedIn profiles. The data contain rich information on workers' employment history and job characteristics at a monthly frequency. Variables include job location (state, city, MSA), seniority levels, detailed occupational and industry codes (SOC and NAICS), firm identifiers with a crosswalk to Compustat as well as detailed job titles and job descriptions. Access is granted via the Wharton Research Data Services website. Revelio Labs and similar resume data have been shown to be a good representation of the U.S. high-skill, professional labor market (Babina et al., 2024; Hampole et al., 2025; Schubert et al., 2025) and consistent with aggregate employment and occupational statistics from CPS and ACS (Tambe, 2025). The data are particularly well suited to study the professional digital intensive workforce that is the subject of study in this section.[11]

2.2 Identifying AI adopters

I classify firms as AI adopters and non-adopters for each year in the sample. This is done in two steps and follows a large body of literature that uses labor market statistics to measure AI adoption (Abis & Veldkamp, 2023; Acemoglu et al., 2022b; Babina et al., 2024; Hampole et al., 2025; Kalyani et al., 2025).

[10] Those are the tech, professional services and high-growth manufacturing industries, i.e. sectors most exposed to AI. Those are also workers who are most likely to be complementary to AI technologies and not substituted by them, at least in the first wave of AI. Note that they may be both complementary to in-house use of AI technologies, but also to integrating outsourced AI tasks into the business, especially for larger firms.

[11] An advantage of resume data is that they contain information about the technologies and tasks used within firms. In contrast, job postings data measure technology demand, but do not capture whether vacancies are filled or filled with workers possessing the demanded skills.

The first step utilizes the rich raw text of the self-reported job titles and descriptions by Revelio Labs to categorize approximately one million workers in the data as using AI, based on technology-related keywords. In a second step, I use the firm identifiers provided by Revelio Labs to map AI workers to individual private firms and Compustat firms. Firms are then classified as an adopter at time t if they employed an AI worker in year t or before.[12] The set of all adopters at time t is denoted by $\mathcal{A}_t$. A firm k at t is referred to as a previous adopter or previously adopting firm if $k \in \mathcal{A}_{t-1}$. A firm j is referred to as a newly adopting firm if $j \in \mathcal{A}_t$ but $j \notin \mathcal{A}_{t-1}$.

Overall, there are three ways in which a firm fills an AI position. They can retrain the existing workforce, they can poach skilled workers from other firms and thereby access the knowledge of third parties, and they can hire out of university. To avoid mechanical issues, some specifications exclude workers used to identify the adoption year of firms when computing labor flows.

Measuring propensity to hire from previous AI adopters. What is the propensity and direction of worker flows from AI utilizing firms to other firms in the economy? I specify the following Poisson model for job-to-job movers. Let $E_{ickj\ell t}$ denote an indicator which takes the value one if a worker i belonging to a group c experiences a job-to-job move from firm k to j in location ℓ at time t. Let $D_{ikt} = \mathbb{1}\{k \in \mathcal{A}_{t-1}\}$ denote the indicator that takes the value one if i's origin firm k belongs to the set of adopters at time $t-1$. The probability of a move for worker i to a firm j is specified as

$$E_{ickj\ell t} = \exp\left\{\alpha_{cj\ell t} + \gamma_{cj\ell t} D_{ikt}\right\} \varepsilon_{ikj\ell t}, \qquad i \in c, \tag{1}$$

with $\mathbb{E}\left[\varepsilon_{ikj\ell t} \mid D_{ikt}, c, j, \ell, t\right] = 1$. The term $\alpha_{cj\ell t}$ is a firm-location-job-mover specific effect that captures the time-varying natural propensity of firm j to absorb workers in set c. This defines the baseline hiring probabilities that will be the reference for constructing the excess propensities.

I refer to $\gamma_{cj\ell t}$ as the *excess probability* of a worker moving from an AI utilizing firm to firm j.[13] It

[12] One concern may be that AI hiring does not necessarily correspond to AI usage. Hampole et al. (2025) document that AI hiring in Revelio Labs correlates strongly with other, more common measures of technology use such as survey-reported AI utilization in the Census Business Trends and Outlook survey (BTOS) and patenting in AI using the USPTO AI patent database.

[13] This is with some abuse, as strictly $e^{\gamma_{cj\ell t}}$ corresponds to the excess probability. However, $\gamma_{cj\ell t}$ maps directly to continuous-time transition rates which will be important in the quantitative model.

is defined relative to a baseline probability $\alpha_{cj\ell t}$ of absorbing a worker from a non-AI utilizing firm. The excess probability nets out factors that affect a firm's hiring uniformly. For example, if a firm is expanding and scales up all forms of hiring by a constant factor, then the excess probability is unaffected. Further, the focus on the conditional probability of a worker moving controls for time-varying changes in the size of pools of AI adopters. Formally, I estimate (1) as the sample equivalent of the log-odds ratio, that is,

$$\gamma_{cj\ell t} \equiv \log\left(\frac{\Pr\left[E_{ickj\ell t} \mid D_{ikt} = 1, i \in c\right]}{\Pr\left[E_{ickj\ell t} \mid D_{ikt} = 0, i \in c\right]}\right).$$

In a second step, I estimate[14]

$$\hat{\gamma}_{cj\ell t+h} = \beta_h \mathbb{1}\{Adoption_{jt}\} + \sum_{l=1}^{2} X_{jt-l}\delta_l + \alpha_\ell + \alpha_n \times \alpha_{t+h} + \nu_{cj\ell t+h}, \tag{2}$$

where $\mathbb{1}\{Adoption_{jt}\}$ takes the value one if firm j is classified as newly adopting AI at time t, α_ℓ are MSA fixed effects, $\alpha_n \times \alpha_{t+h}$ are 3 digit NAICS industry-time fixed effects and X are pre-adoption controls that include log sales per workers deflated by NIPA prices, log professional employment, and professional share. The coefficients β_h then capture how relative hiring propensities change with time since AI adoption h.

Intuitively, this regression traces how hiring flows are associated with the distance to the adoption year. For example, for a firm that relies exclusively on internal training or hiring out of university, we expect β_h to not significantly change over horizon h; i.e. there not to be a change in poaching behavior. In contrast, for firms which use poaching to access the talent pool of AI utilizing firms in order to adopt, we expect β_h to increase around the time of adoption.

2.3 Hiring from previous adopters increases around time of adoption

The baseline specification defines the set c as all professional workers with an employer-to-employer (E-E) move in metropolitan MSAs in the U.S. who move from some firm k to a publicly traded firm j. Figure 1 plots the coefficients β_h for the distance to adoption year $h \in [-4, 5]$. There is a discontinuous jump in

[14] In Appendix C.4, I discuss the variation exploited in the regression as seen through the lens of the quantitative model in the subsequent section.

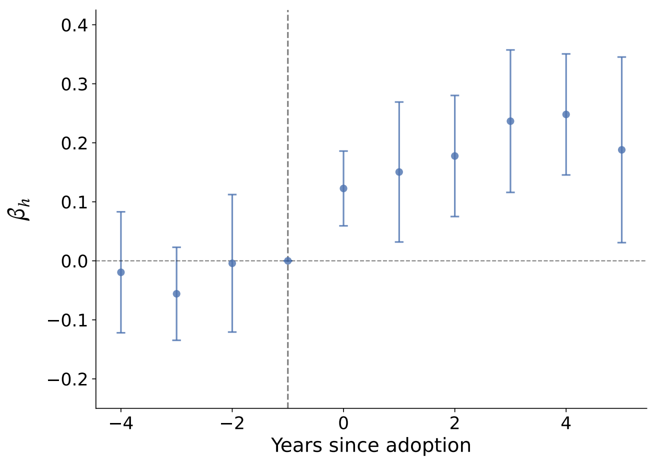


Figure 1: Excess propensity to hire from a previous adopter.

Notes: This graph plots the coefficient estimates β_h from (2) and corresponds to how excess probabilities $\gamma_{cj\ell t+h}$ change with distance to the adoption year t. The coefficients are estimated on subsample of E-E movers at annual frequency. The regression controls for industry-year and location fixed effects as well as pre-adoption variables related to selection into adoption (e.g., sales per worker). The sample period is 2010-2024. Destination firms in the sample are Compustat firms. Origin firms in the sample are the universe of firms in Revelio Labs. Standard errors are clustered at the firm-year level.

excess probabilities at the time of adoption of about 15% and a peak response of approximately 25%.

The discontinuity suggests that this correlation is specific to the timing of adoption events and does not capture smooth underlying trends, while the increase in excess probabilities is persistent. Taken together, the descriptive evidence is consistent with the view that worker flows are important at the time of AI adoption and that AI adoption is accompanied by discrete changes in hiring practices to source outside talent. Tapping into different talent pools seems to be a salient feature of the data even after controlling for productivity and factors at the industry-time level.

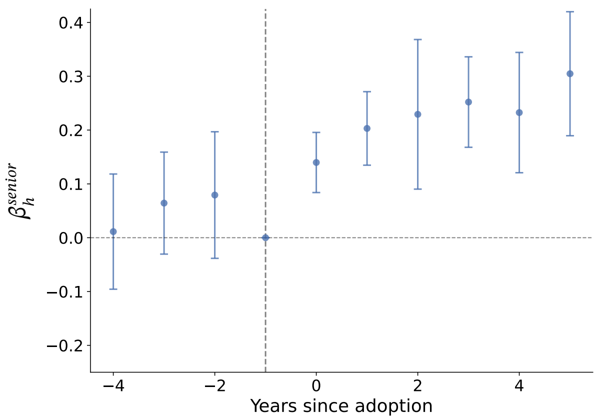
A baseline estimate of 0.15 implies a 15% increase in the log-odds ratio, translating to approximately $e^{0.15} - 1 \approx 0.162$ or a 16.2% increase in the relative probability that firms hire from a previous adopter compared to their baseline hiring patterns. The magnitude of the baseline estimates compare to an interquartile range of excess probabilities in the data of about 0.8. The estimates are therefore meaningful, but not unreasonably large. Importantly, these estimates capture changes across the entire professional workforce, not just inventors as previously studied in the literature. These estimates are robust to a series of variations explored in Appendix C.5.

One concern is that our adoption measure, employing AI workers, and our outcome, hiring from AI utilizing firms, both involve labor market measures, potentially creating a mechanical relationship. I address this in several ways. First, I re-estimate the above specifications excluding the first AI workers in a firm. This ensures that the workers used to classify adoption events are not simultaneously used to measure worker flows (see Appendix C.5.1). Figure C.4 shows that estimates are both qualitatively and quantitatively robust to this sample restriction. Second, insofar a mechanical relationship exists it would be concentrated at horizon $h = 0$. Effects are, however, strongly persistent and approximately constant. Third, our adoption measure is not determined by *cross-firm* labor flows necessarily. In total there are three distinct channels: (i) promotion of existing employees to AI roles (ii) hiring from universities and (iii) poaching from other firms.

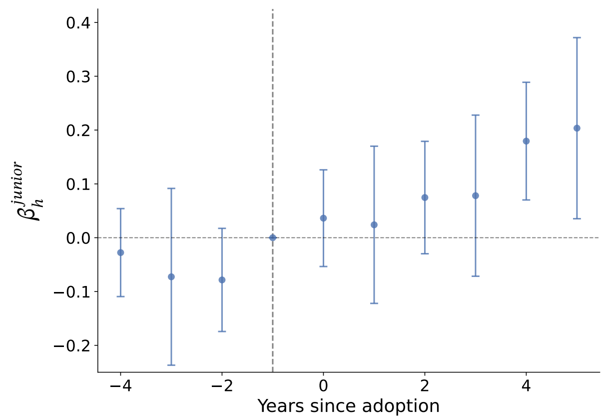
Hiring increases concentrated for senior workers. Next, I examine for which types of workers the associations are strongest. I estimate (2) for two subsamples in the data: (i) letting c be the set of all *senior* workers defined as workers with job titles corresponding to seniority levels of managers, directors and executives and (ii) letting c be the set of all *junior* workers. Estimates of coefficients are reported in Figure 2. In the data, senior hires jumps up immediately at the time of adoption while the coefficients for junior hires only become significant after year 4. Again, estimates are robust to excluding the first AI workers within a firm (see Figure C.6).

This is consistent with the view that previous experience accumulated over time and practical application of AI technology is most important in the early phases of adoption and more difficult to access through junior hires. I view this as indicative of the view that firms use poaching to adopt a technology and not simply to access technical skills after successfully adopting AI. But as AI use matures inside a firm, technical skills may become the primary constraining factor and junior hires increase.[15]

[15]These results are consistent with the view that AI adoption leads to teams becoming more junior (Babina et al., 2024), but that senior workers are necessary to form such AI teams at the start. This provides a different perspective on the recently documented changes in senior and junior employments of GenAI adopters in the U.S. (Brynjolfsson et al., 2025; Lichtinger & Hosseini Maasoum, 2025): these trends may reflect the demand of practical implementation knowledge around GenAI adoption.



(a) Senior workers.



(b) Junior worker.

Figure 2: Excess propensity to hire from a previous adopter by seniority status.

Notes: This graph plots the coefficient estimates β_h^c from (2) and corresponds to how excess probabilities $\gamma_{cj\ell t+h}$ change with distance to the adoption year t, separately for two subgroups c: senior workers (e.g. managers and directors) and junior workers (e.g. analysts). The left panel shows coefficients for the senior subsample. The right panel shows coefficients for the junior subsample. Coefficients are estimated on a subsample of E-E movers at annual frequency. The regression controls for industry-year and location fixed effects as well as pre-adoption variables related to selection into adoption (e.g. sales per worker). The sample period is 2010-2024. Destination firms in the sample are Compustat firms. Origin firms in the sample is the universe of firms in Revelio Labs. Standard errors are clustered at the firm-year level.

3 Model

How does a radical new technology diffuse across an economy, and how does this process differ from incremental productivity improvements? In this section, I propose an equilibrium model that studies the dynamics of an economy after an introduction of a new technology in a setting of mutual firm and worker learning with endogenous technology adoption and embedded in in a standard job ladder model with on-the-job search and knowledge spillovers.

Environment. Time is continuous and runs forever. At any time $t \in \mathbb{R}$, the economy is endowed with a set $\mathcal{T}_t$ of technology vintages. Prior to $t = 0$, there exists a single technology vintage, so $\mathcal{T}_t = \{0\}$. At $t = 0$, a new vintage becomes available exogenously and thereafter $\mathcal{T}_t = \{0, 1\}$ for all $t \geq 0$. The arrival of the new technology is unanticipated prior to $t = 0$. There are no other aggregate shocks.

The economy is populated by two types of agents: a unit mass of workers and a mass M of entrepreneurs. Agents are risk neutral and have perfect foresight over the equilibrium paths of aggregate variables.

3.1 Workers

Workers are infinitely lived and discount the future at rate ρ. Their intertemporal utility is linear and workers maximize the net present value of discounted utility over consumption. At any time t, a worker is indexed by three states: (i) the most recent technology vintage in which the worker has experience $\tau \in \{0, 1\}$, (ii) the vintage-specific human capital $h \in \mathbb{R}^+$, and (iii) their employment status. If a worker is employed, they are further indexed by the firm that they are matched with.

Employed workers inelastically supply one unit of labor to entrepreneurs and receive an endogenous wage w_t which is determined following the sequential auction bargaining protocol by Postel-Vinay and Robin (2002) according to which entrepreneurs Bertrand compete for the worker. Unemployed workers receive unemployment benefits bh. As described in detail below, employed workers accumulate human capital specific to their vintage through on-the-job learning. When unemployed, their human capital depreciates at rate α_u.

3.2 Entrepreneurs

Entrepreneurs can either be matched with a single worker or operate solo. Matched entrepreneurs produce output and decide whether to upgrade their technology vintage. Solo entrepreneurs produce output and post vacancies to hire a worker either out of unemployment or poach an employed worker, but cannot adopt a new vintage without a worker who has experience with the new technology. Both types of entrepreneurs discount the future at the real interest rate r_t.

3.2.1 Matched entrepreneur

The state of a matched entrepreneur is given by the tuple (z, h, τ), where $z \in \mathbb{R}^+$ is the entrepreneur's idiosyncratic, vintage-specific productivity, h the worker's human capital, and $\tau \in \{0, 1\}$ the technology vintage that is used in operation. The production process combines z, h and τ into net output $\pi(z, h \mid \tau) = A^\tau \left(z^\zeta + h^\zeta\right)^{1/\zeta} - \tau k_f$. The production function features complementarities between a worker's human capital and an entrepreneur's productivity whenever it is supermodular, i.e. $\zeta < 1$, regardless of technology vintage—which is satisfied in the calibrated model. The new technology vintage, i.e. $\tau = 1$, modifies the production process in two ways. First, it scales output by a factor $A > 1$. Second, it introduces a per-period fixed cost $k_f > 0$. This captures the idea that modern technologies which heavily rely on intangible investments tend to lower marginal costs while increasing fixed costs in production (De Ridder, 2024).[16] A match is exogenously separated at rate $\varkappa$ at which point the worker becomes unemployed and the entrepreneur solo.

Learning dynamics. During a match, firm and worker knowledge evolve endogenously as a two-sided process. Similar to the "learning from coworker" literature, a worker learns from the entrepreneur at a rate proportional to the knowledge gap, $\alpha [z - h]^+$, where $x^+ = \max\{x, 0\}$. At the same time, worker knowledge gets embodied in the firm at a rate $\alpha [h - z]^+$. For example, workers may write manuals transcribing their knowledge or pass on tacit knowledge about ways in which a technology can or cannot be operationalized. The simple process outlined below is a stand-in for all such forms of within match knowledge spillovers. Agents only learn from a more knowledgeable peer and cannot unlearn through low quality matches. This assumption is motivated by empirical evidence from the learning from coworkers literature (Herkenhoff et al., 2024; Jarosch et al., 2021). In addition, firm productivity is subject to exogenous idiosyncratic shocks, which capture all other shocks to a match's profitability such as management skills, business opportunities, idiosyncratic demand shocks, etc..

Taken together, the law of motions of z and h within a match are given by

$$\begin{aligned} dz_t &= \alpha [h_t - z_t]^+ dt + \vartheta (\mu - \log z_t) z_t dt + \sigma z_t dB_t \\ dh_t &= \alpha [z_t - h_t]^+ . \end{aligned} \tag{3}$$

The first term of the first line captures the adjustment due to the knowledge spillovers, while

[16] In the present model, the fixed cost also ensures that it is the most productive entrepreneurs that first adopt the new technology, consistent with the data (Hampole et al., 2025).

the second term captures the mean-reverting nature of exogenous shocks and dB_t denotes a standard Brownian motion. Combined, the stochastic process for the entrepreneur's productivity is a geometric Ornstein-Uhlenbeck process that incorporates knowledge spillover. The second line captures the on the job learning by workers. If the entrepreneur is more knowledgeable, the worker's human capital catches up at a rate proportional to the knowledge gap. There are no shocks to worker's human capital directly, though human capital inherits the stochastic nature through its coupling with the entrepreneur's productivity.

The process in (3) defines an asymmetric learning process: While the upside of volatility from positive productivity shocks is captured, the learning process protects against the downside risk of negative shocks. That is, a series of sufficiently many positive shocks to the entrepreneur's productivity induces positive human capital growth, whereas a series of sufficiently negative shocks leading to $z < h$ does not reduce human capital, but creates the possibility for the entrepreneur to catch up to the worker's human capital through the mutual learning process. This process then delivers positive trend growth while a match persists (see Appendix A.7). This captures the process of innovation or learning about a technology as in Atkeson and Kehoe (2007) in a reduced form way. Overall, the learning process generates returns to match duration.

Technology adoption. Beyond incremental learning within an existing technology vintage, worker-entrepreneur pairs can undertake discrete technology upgrades. At the start of each period, the match jointly decides whether to adopt a new technology vintage. Upgrading to a new technology vintage entails the loss of vintage-specific expertise of both entrepreneur and workers: there is an incomplete and stochastic transfer of existing knowledge between old and newer technology vintages. Specifically, upon choosing to use technology vintage τ with initial knowledge states (z_{t-}, τ_f) and (h_{t-}, τ_w) for entrepreneur and worker, respectively, their productivity and human capital in the new vintage become $z_t = \delta^{\tau(1-\tau_f)} z_{t-}$ and $h_t = \delta^{\tau(1-\tau_w)} h_{t-}$.[17] The formulation captures that only agents without prior technology experience, i.e. $\tau_i = 0$ for $i \in \{w, f\}$, are subject to the partial transferability process. The knowledge transfer para-

[17] This is written in parsimonious form. If the match decides to adopt, then $\tau = 1$. If the worker has prior experience in the new technology, then $\tau_w = 1$. If the firm has prior experience in the new technology, then $\tau_f = 1$. Taken together, limited transferability only arises whenever (i) the match is upgrading to $\tau = 1$ and (ii) one or both agents have no prior experience, i.e. $\tau_i = 0, i \in \{w, f\}$.

meter δ follows the two-state process[18]

$$\delta = \begin{cases} \overline{\delta} & \text{w.p. } \mathcal{P}(\tau_f, \tau_w), \\ \underline{\delta} & \text{otherwise,} \end{cases} \tag{4}$$

where $\mathcal{P}(\cdot,\cdot)$ denotes the probability of drawing the high-probability state and may differ depending on prior technology experience, thereby allowing for knowledge spillovers.

The mapping of knowledge between vintages has the following key properties. First, knowledge is fully transferable for the agent that has used the technology previously. For example, suppose the worker has prior experience (i.e. $\tau_w = 1$) with the technology and the entrepreneur does not (i.e. $\tau_f = 0$), then knowledge with the new technology ($\tau = 1$) is $z_t = \delta z_{t-}$ and $h_t = h_{t-}$. In this sense, it is cheaper for an entrepreneur to adopt the new vintage when hiring a worker with prior experience. Second, there are direct knowledge spillovers from adopters to non-adopters captured by the stochastic process and the probability mapping $\mathcal{P}(\tau_f, \tau_w)$ if $\min\{\mathcal{P}(1,0), \mathcal{P}(0,1)\} > \mathcal{P}(0,0)$. This captures the fact that experience with the new technology can help best apply the old knowledge to the new tasks and therefore improves the chances of greater transferability.[19] At last, knowledge is fully transferable when downgrading to an older technology vintage (i.e. $\tau = 0$).[20]

3.2.2 Solo entrepreneur

The state of a solo entrepreneur is given by the tuple (z, τ), where z and τ are defined analogously to above. Production occurs without a worker, so net output is $\pi(z, 0 \mid \tau) = A^\tau z - \tau k_f$. Productivity is determined exogenously by the mean-reverting Ornstein-Uhlenbeck process. There is no opportunity to upgrade the technology vintage without a worker. Solo entrepreneurs face an additional exogenous risk of exiting d in which case they get a value of zero and are replaced by an identical solo entrepreneur. This increases

[18]Instead of assuming a two-state stochastic process, we could also assume that the transferability parameter depends on τ_f and τ_w, i.e. $\delta(\tau_f, \tau_w)$. This delivers qualitatively the same results and would be slightly more parsimonious.

[19]An alternative interpretation is that prior experience helps with avoiding crucial mistakes in the beginning.

[20]Downgrading may occur whenever worker and entrepreneur do not have the same technology state, i.e. $\tau_f \neq \tau_w$. However, I assume that if worker and entrepreneur both have used the new technology vintage, i.e. $\tau_f = \tau_w = 1$, then they do not downgrade their technology, i.e. it is an absorbing state. In the calibrated model this assumption is non-restrictive.

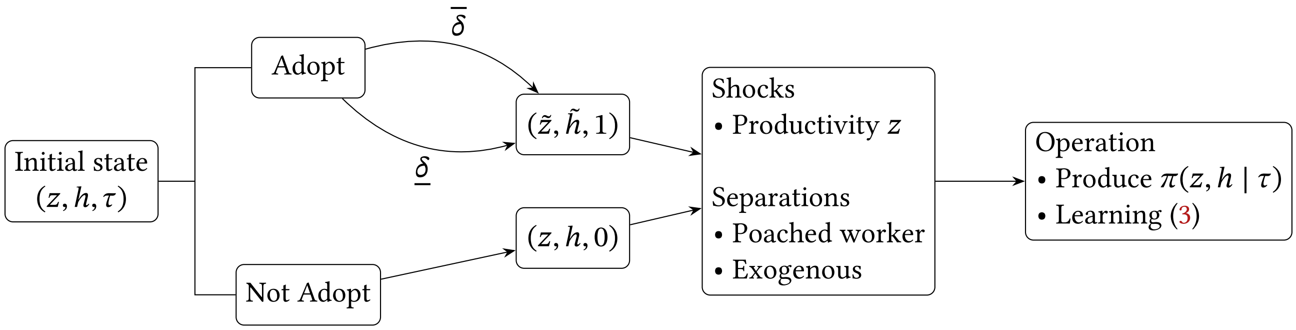


Figure 3: Decision tree for hiring and technology adoption by matched entrepreneur.

Notes: This timeline shows the sequence of events for a matched entrepreneur. The period is entered with an initial state (z, h, τ). The match first decides whether to adopt the new technology vintage. Upon adopting the match is subject to the stochastic transferability parameter, δ. The new state variables realize and the rest of the period unfolds: productivity shocks, exogenous and endogenous separations alongside production and learning.

the effective discount rate of solo entrepreneurs to $r_t + d$ and captures that letting go of the worker incurs additional risk.

A solo entrepreneur posts vacancies v at a convex cost $c(v) = \bar{c}/(1+\gamma)v^{1+\gamma}$ with $\gamma > 0$, which generates meetings with a worker at rate $q_t v$. Entrepreneurs take the aggregate contact rate q_t as given, but choose the number of vacancies v to maximize expected profits. Upon meeting, entrepreneur and worker enter the sequential auction bargaining protocol of Postel-Vinay and Robin (2002). If a new match is formed, entrepreneur and worker optimally choose the technology vintage based on their current knowledge (z, τ_f) and (h, τ_w). The transferability of knowledge from old to new vintage takes the same form as in (4). The sequence of events is summarized in Figure 4.

3.3 Matching technology and bargaining

Solo entrepreneurs and workers meet in a frictional labor market. Search is random. Meetings are mediated through a constant returns to scale matching function $m(\mathbb{X}_t, \mathbb{V}_t) = \chi \mathbb{X}_t^{1-\beta} \mathbb{V}_t^{\beta}$, where $\mathbb{X}_t = u_t + \xi(1-u_t)$ denotes the effective pool of job seekers and $\mathbb{V}_t$ the aggregate vacancies. The aggregate unemployment rate is denoted by u_t and ξ the relative search intensity of employed workers. Labor market tightness is denoted by $\theta_t = \mathbb{V}_t/\mathbb{X}_t$. The contact rates for workers and solo entrepreneurs are given by $p(\theta_t) = m(\mathbb{X}_t, \mathbb{V}_t)/\mathbb{X}_t =$

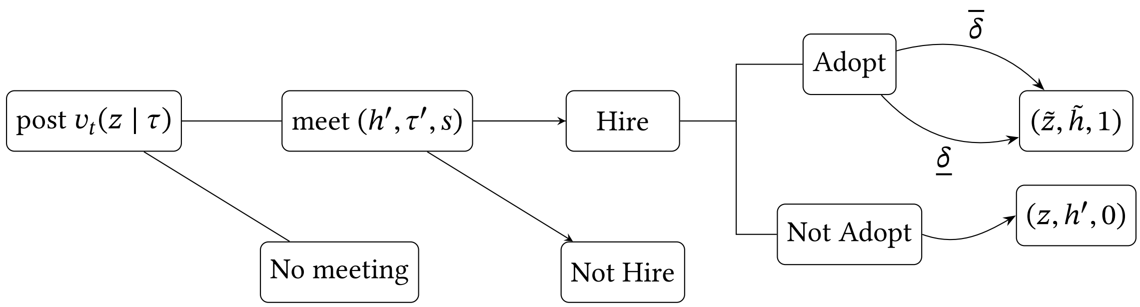


Figure 4: Decision tree for hiring and technology adoption.

Notes: This timelines shows the sequence of events of a solo entrepreneur posting vacancies and searching for a worker. A worker's employment state is denoted by $s \in \{U, E(z', \tau')\}$ where employment state E is indexed by an entrepreneur (z', τ') who the worker is matched to—which is important for the worker's outside option in bargaining. If a worker is hired, the new match decides on which technology vintage to use subject to limited transferability of prior knowledge. At the end, worker and entrepreneur agree on the same technology vintage.

$\chi\theta_t^\beta$ and $q(\theta_t) = p(\theta_t)/\theta_t$, respectively. An unemployed worker is contacted at a rate $q(\theta_t)$ per vacancy posted while an employed worker gets contacted at a rate $\xi q(\theta_t)$. Likewise, an unemployed worker meets a solo entrepreneur at a rate $p(\theta_t)$, while an employed worker meets a solo entrepreneur at a rate $\xi p(\theta_t)$.

Match surplus is divided between the entrepreneur and the worker following the sequential auction framework developed by Postel-Vinay and Robin (2002). Entrepreneurs Bertrand-compete through a sequence of wage bidding contests. That is, when an employed worker receives an outside offer, the incumbent employer has the option to counter-offer, and the winner is the firm that has the higher match surplus with the worker consistent with the allocation in a first price ascending auction. The winning bidder pays just enough to leave the losing firm indifferent between matching the offer and retaining and losing the worker. When an unemployed worker receives an outside offer, the entrepreneur hires the worker if the match generates a positive surplus, in which case they promise a value to the worker that leaves them indifferent between working and remaining unemployed. Wages adjust to implement the allocation, which is feasible since utility is transferable. Overall, the bargaining protocol implies that worker mobility is bilaterally efficient.

3.4 Recursive representation

Having described the economy's environment, this section details the value functions that determine the *allocative* decisions in the economy and formalizes the environment described in the preceding section. The sequential auction protocol together with transferable utility implies that allocative decisions do not depend on the split of the value of a match between worker and entrepreneur, but only on its total value and the combine state (z, h, τ).[21] As utility is linear over consumption, the real interest rate is pinned down by the discount rate, i.e. $r_t = \rho$ for all t.

Let the total value of a match between an entrepreneur and a worker be denoted by $S_t(z, h \mid \tau)$, the value of a solo entrepreneur by $V_t(z \mid \tau)$, and the value of an unemployed worker by $U_t(h \mid \tau)$. The total value of a job, $S_t(z, h, \tau)$, satisfies the following HJB[22]

$$\begin{aligned}\rho S_t(z,h|\tau) - \partial_t S_t(z,h \mid \tau) = &\underbrace{\pi(z,h|\tau)}_{\text{flow surplus}} + \underbrace{\partial_z S_t(z,h|\tau)\alpha\,[h-z]^+ + \partial_h S_t(z,h|\tau)\alpha\,[z-h]^+}_{\text{learning gain}} \\ &+ \underbrace{\partial_z S_t(z,h|\tau)\vartheta(\mu - \log z)z + \frac{\sigma^2 z^2}{2}\partial_{zz} S_t(z,h|\tau)}_{\text{exogenous shocks}} \\ &+ \underbrace{\kappa\,(V_t(z|\tau) + U_t(h|\tau) - S_t(z,h|\tau))}_{\text{separation shock}},\end{aligned} \tag{5}$$

subject to the boundary conditions for adoption and endogenous separation

$$S_t(z,h|\tau) \geq S_t^A(z,h \mid \tau,\tau), \qquad \text{and} \qquad S_t(z,h|\tau) \geq V_t(z|\tau) + U_t(h|\tau),$$

where the value of adopting is given by $S_t^A(z, h \mid \tau_f, \tau_w) \equiv \mathbb{E}\left[S_t(\tilde{z}, \tilde{h} \mid 1) \mid z, h, \tau_f, \tau_w\right]$, where $\tilde{z} = \delta^{1-\tau_f} z$ and $\tilde{h} = \delta^{1-\tau_w} h$ and δ follows the stochastic process defined in (4).

Let the endogenous cumulative distribution function of active matches, and the unemployed be

[21] Technically, we must also assume that adoption decisions are made jointly by worker and entrepreneur in order to maximize the total value. Bilal et al. (2022) show how this can be rationalized as the outcome of a bargaining process. Intuitively, if there are actions that could increase the total value of a job there should exist a bargaining protocol that apportions the surplus.

[22] For notational convenience, we define the differential operators $\partial_z = \frac{\partial}{\partial z}$, $\partial_{zz} = \frac{\partial^2}{\partial z^2}$ and $\partial_h = \frac{\partial}{\partial h}$.

given by $G_t^e(z, h \mid \tau)$ and $G_t^u(h \mid \tau)$, respectively. Define the value of an entrepreneur with states (z, τ) hiring a worker with states (h', τ') to be $S_t^H(z, h \mid \tau_f, \tau_w) = \max\{S_t^A(z, h \mid \tau_f, \tau_w),\ S_t(z, h \mid 0)\}$, taking into account the optimal choice of the technology vintage. Then, the value of a solo entrepreneur is

$$
\begin{aligned}
(\rho + d)V_t(z|\tau) - \partial_t V_t(z|\tau) = &\underbrace{\pi(z, 0|\tau)}_{\text{flow profit}} + \underbrace{\partial_z V_t(z|\tau)\vartheta(\mu - \log z)\, z + \frac{\sigma^2 z^2}{2}\partial_{zz}V_t(z|\tau)}_{\text{exogenous shocks}} \\
&\underbrace{+ \max_v \left\{ vq(\theta_t)\frac{1}{\mathbb{X}_t}\left[\xi \sum_{\tau' \in \mathcal{T}_t} \int \Delta_t^{EE}(z', h', z|\tau', \tau)^+ dG_t^e(z', h'|\tau') + \sum_{\tau' \in \mathcal{T}_t} \int \Delta_t^{UE}(h', z|\tau', \tau)^+ dG_t^u(h'|\tau') \right] - c(v) \right\}}_{\text{expected gain from posting vacancies}},
\end{aligned} \tag{6}
$$

where the expected gain from posting vacancies aggregates over the individual gains of potential matches

$$
\begin{aligned}
\Delta_t^{EE}(z', h', z \mid \tau', \tau) &= S_t^H(z, h' \mid \tau, \tau') - V_t(z|\tau) - \left(S_t(z', h' \mid \tau') - V_t(z' \mid \tau')\right) \\
\Delta_t^{UE}(h', z \mid \tau', \tau) &= S_t^H(z, h' \mid \tau, \tau') - V_t(z|\tau) - U_t(h|\tau').
\end{aligned} \tag{7}
$$

The value of unemployment solves

$$
\rho U_t(h \mid \tau) - \partial_t U_t(h \mid \tau) = bh - \partial_h U_t(h \mid \tau)\alpha_u h. \tag{8}
$$

Given the boundary condition $U_t(0 \mid \tau) = 0$, the unique solution is $U_t(h \mid \tau) = bh/(\rho + \alpha_u)$.

Comment. Value functions such as $S_t(z, h \mid \tau)$ are time-varying because they depend on the aggregate state through the option value of becoming a solo entrepreneur. The key aggregate variables are labor market tightness θ_t and the endogenous distribution over matches and unemployed G_t^e and G_t^u, which determine the expected gains from posting vacancies. The value of a match inherits this time dependency through the option value of becoming a solo entrepreneur either through endogenous or exogenous separations. It follows that in a model without exogenous separations, i.e. $\kappa/\rho \to 0$, and no endogenous separations the total value of a match is constant over the transition towards the new ergodic steady-state. Conversely, in the full model all time-variation in policy functions over the transition stem from changes in the option value.

3.5 Policy functions

Agents in the economy make a total of five decisions: number of vacancies posted, v_t, discrete hiring choices, $\mathbb{1}^{EE}$, and $\mathbb{1}^{UE}$, discrete adoption choice, $\mathbb{1}^{A}$, and discrete match separation choice, $\mathbb{1}^{X}$.

Decisions of solo entrepreneur. First, solo entrepreneurs choose how many vacancies to post, taking as given the aggregate matching rate, $q(\theta_t)$, and the distribution over potential matches dG_t^e and dG_t^u. The first-order condition of (6) yields the optimal vacancy posting policy

$$\bar{c}v_t(z \mid \tau)^{\gamma} = q(\theta_t)\frac{1}{\mathbb{X}_t}\left[\xi \sum_{\tau' \in \mathcal{T}_t} \int \Delta_t^{EE}(z', h', z|\tau', \tau)^{+} dG_t^e(z', h'|\tau') + \sum_{\tau \in \mathcal{T}_t} \int \Delta_t^{UE}(h', z|\tau', \tau)^{+} dG_t^u(h'|\tau')\right].$$

The left-hand side denotes the marginal cost of posting a vacancy. The right-hand side is the expected surplus gain from posting an additional vacancy, which is determined by the offer distributions of Δ^{EE} and Δ^{UE} defined in (7). Vacancy posting is decreasing in labor market tightness, θ_t, and increasing in the expected surplus gain averaged over the population distributions, G_t^e and G_t^u. Matched workers tend to have a higher human capital which generates larger expected match values, but poaching then is more expensive than hiring out of unemployment. By contrast, unemployed workers have a lower outside option and entrepreneurs can extract a larger fraction of the surplus. A greater availability of workers with knowledge in the new technology vintage tend to increase the expected value of new matches.

Hiring decision. There are three primary forces that determine the hiring policies: (i) production gains and complementarities in z and h, (ii) learning gains governed by the learning rate α as well as the scope of learning $|z - h|$, and (iii) the value of outside options V_t and U_t. The option to adopt a new technology vintage creates opposing forces. While it strengthens non-adopters' incentives to poach experienced workers and upgrade to the new vintage, it also intensifies previous adopters' retention incentives, as worker productivity increases with the new technology. Details are provided in Appendix A.3, including a decomposition of these factors.

Conditional on meeting with an employed or unemployed worker the solo entrepreneur next engages in the sequential auction to bid for the worker. As discussed, a new match is formed whenever it is bilaterally efficient to do so. From (7) this is determined by Δ_t^{EE} and Δ_t^{UE}. The hiring policies are therefore given by the indicator functions

$$\mathbb{1}^{EE}\left(z', h', z \mid \tau_w, \tau_f; t\right) = \mathbb{1}\left\{\Delta_t^{EE}(z', h', z \mid \tau_w, \tau_f) > 0\right\}$$
$$\mathbb{1}^{UE}\left(h', z \mid \tau_w, \tau_f; t\right) = \mathbb{1}\left\{\Delta_t^{UE}(h', z \mid \tau_w, \tau_f) > 0\right\}.$$

Adoption decision. Once a match is formed, the entrepreneur-worker pair decides which technology is optimal to use during operation. The match uses the new technology vintage whenever

$$\mathbb{1}^{A}\left(z, h \mid \tau_f, \tau_w; t\right) = \mathbb{1}\left\{S_t^A(z, h \mid \tau_f, \tau_w) \geq S_t(z, h \mid 0)\right\}.$$

The combination of per-period fixed cost and greater productivity with the modern technology implies a cut-off productivity $\underline{z}_t(h)$ such that for all $z \geq \underline{z}_t(h)$ adoption occurs and no adoption otherwise.[23] This is a quantitative result in line with empirical results around AI adoption, documenting that larger, high sales, and more productive firms are more likely to adopt (Hampole et al., 2025).

Separation decision of matched entrepreneur. At the beginning of each period matched entrepreneurs decide whether to terminate the match $\mathbb{1}^X(z, h \mid \tau; t) = \mathbb{1}^X\{S_t(z, h \mid \tau) \leq V_t(z \mid \tau) + U_t(h \mid \tau)\}$.

3.6 Equilibrium

The economy before the arrival of the new technology vintage, i.e. $t < 0$, is in its ergodic steady-state. That stationary equilibrium is defined as follows.

Definition 1 (Stationary equilibrium)**.** Given $\mathcal{T}_t = \{0\}\ \forall t < 0$, the initial stationary equilibrium prior to $t = 0$ is defined as the set of value functions $\{S_{ss}, V_{ss}, U_{ss}\}$, distribution functions $\{G_{ss}^e, G_{ss}^v, G_{ss}^u\}$, policy functions $\{v_{ss}, \mathbb{1}_{ss}^{EE}, \mathbb{1}_{ss}^{UE}, \mathbb{1}_{ss}^{A}, \mathbb{1}_{ss}^{X}\}$, labor market tightness θ_{ss} such that (5), (6), and (8) hold, labor market tightness is constant and consistent with aggregate vacancies $\mathbb{V}_{ss} = \int v_{ss}(z \mid 0) dG_{ss}^v(z \mid 0)$ and aggregate

[23] Technically, there is an opposite force coming from the assumption on limited transferability of knowledge. For example, lower productivity entrepreneurs have less to lose upon switching to the new technology. In the calibration used in the main exercise this effect is confirmed to not be large and to not dominate.

search input $\mathbb{X}_{ss} = \int dG^u_{ss}(h \mid 0) + \xi \int dG^e_{ss}(z, h \mid 0)$, and distributional dynamics are constant and consistent with policies and worker flows.

After the introduction of the new technology vintage, i.e. $t \geq 0$, the economy transitions to its new steady-state. The equilibrium over the transition is defined as follows.

Definition 2 (Transition path)**.** Given the initial distribution over matches, solo entrepreneurs, and unemployed $\{G^e_{ss}, G^v_{ss}, G^u_{ss}\}$ an equilibrium over the transition after the unanticipated arrival of the new technology vintage is characterized by the full sequence of value functions $\{S_t, V_t, U_t\}$, distribution functions $\{G^e_t, G^v_t, G^u_t\}$, policy functions $\{v_t, \mathbb{1}^{EE}, \mathbb{1}^{UE}, \mathbb{1}^A, \mathbb{1}^X\}$, and labor market tightness $\{\theta_t\}$ such that (5), (6), and (8) hold, labor market tightness is consistent with aggregate vacancies $\mathbb{V}_t = \sum_{\tau \in \mathcal{T}_t} \int v_t(z \mid \tau) dG^v_t(z \mid \tau)$ and aggregate search input $\mathbb{X}_t = \sum_{\tau \in \mathcal{T}_t} \int dG^u_t(h \mid \tau) + \xi \sum_{\tau \in \mathcal{T}_t} \int dG^e_t(z, h \mid \tau)$, and distributional dynamics are consistent with policies and worker flows. A complete characterization of the distributional dynamics using the Kolmogorov Forward Equation is provided in Appendix A.1.

3.7 Properties of the model

This section discusses key definitions and model properties.

Labor mobility. Worker mobility serves three distinct functions. First, reallocation facilitates assortative matching based on human capital and firm productivity, generating output gains through production complementarities since $\zeta < 1$. Second, worker transitions enable knowledge diffusion by generating matches with positive learning gains. Third, labor flows facilitate technology adoption through the transfer of vintage-specific knowledge embodied in workers and firms and reducing the effective cost of adoption.

Knowledge diffusion. This is the act of transfer of knowledge between two agents. I distinguish between two types. *Gradual diffusion* occurs within ongoing employment relationships through incremental learning. *Discrete diffusion* occurs at match formation when technology adoption decisions are made, generating discontinuous changes in knowledge and technology. The former benefits from match duration while the latter occurs at a distinct point in time.

Channels of adoption. Entrepreneurs and workers can adopt the frontier technology vintage either internally or externally. I distinguish between the following channels.

Definition 3 (Adoption by entrepreneurs)**.** Entrepreneurs adopt the new technology through

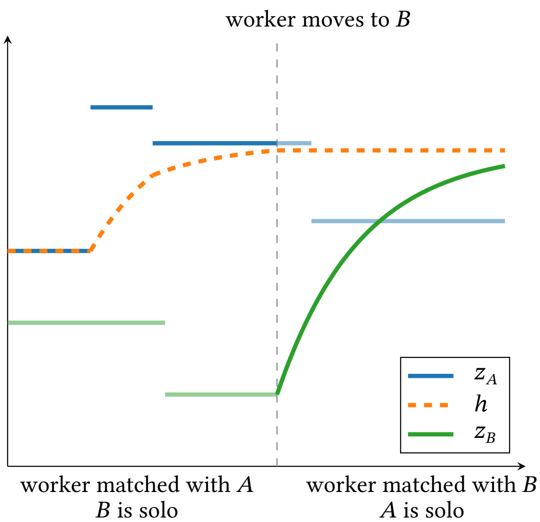
1. *Internal adoption*: Existing matches upgrade to the new vintage subject to the incomplete transferability of their prior knowledge.
2. *Adoption via hiring*: Entrepreneurs without the modern technology hire experienced workers and adopt the new vintage upon forming the new match. This channel is further divided into poaching and hiring out of unemployment.

Definition 4 (Adoption by workers)**.** Workers adopt the new technology through

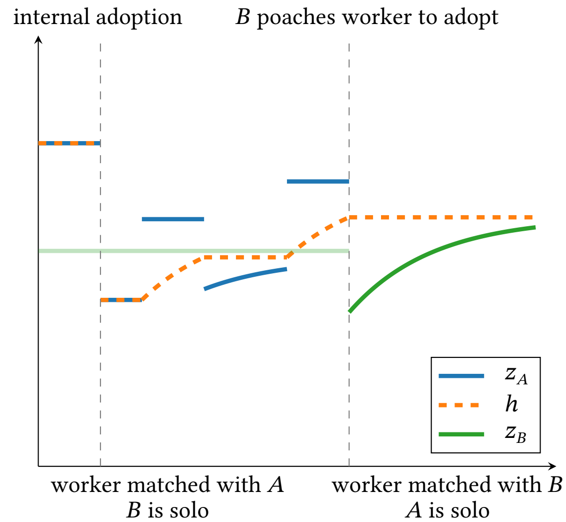
1. *Internal adoption*: Existing matches upgrade to the new vintage subject to the incomplete transferability of their prior knowledge. This coincides with point 1 from Definition 3.
2. *Adoption via training*: Experienced entrepreneurs hire inexperienced workers and use the modern technology in the new match. This channel is further divided into poaching and hiring out of unemployment.

Example paths. Figure 5 plots illustrative example paths and demonstrates key dynamics of the model. In the illustrations, there are two entrepreneurs, A and B, and a single worker.

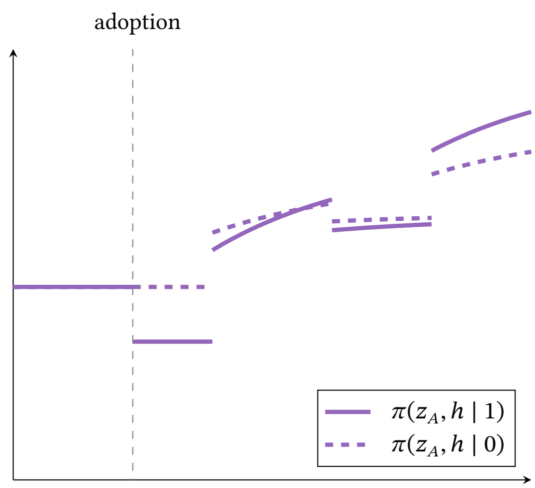
Learning and poaching within vintage. Figure 5a depicts dynamics relating to learning and poaching within a given vintage. Initially, the worker is matched with entrepreneur A while entrepreneur B operates solo. Worker and entrepreneur A have identical knowledge levels z_A and h at the start. At random times, entrepreneur A experiences exogenous shocks to their productivity. A gap between the productivity of the entrepreneur A and the worker's human capital emerges, generating learning gains. The worker then incrementally learns from the entrepreneur and sees their human capital increase. Further, as long as learning gains are present the value of maintaining the match is relatively large. As the gap closes over time, gains from continuing the match fall, making it more likely for the worker to get poached. At a random time, entrepreneur B meets with the worker. If entrepreneur B has sufficiently large learning gains from hiring the worker, the worker moves from A to B even if the latter has a lower productivity. Sub-



(a) Poaching and learning dynamics.



(b) Adoption dynamics.



(c) Output under adoption and non-adoption.

Figure 5: Example paths.

Notes: The top left panel shows productivity paths for two entrepreneurs A and B and the human capital path for a single worker. The worker is initially matched to A. The panel shows the learning dynamics within an existing match and when switching employer. The right panel shows productivity paths and the human capital path. The panel illustrates the limited transferability assumption when upgrading to a new technology vintage and the advantage as a technology laggard (here entrepreneur B) of poaching an experienced worker to adopt. The bottom right panel shows the output paths for a single matched entrepreneur A if they adopt (solid) and if they do not adopt (dashed). The match is subject to identical sequence of shocks, but adoption comes with limited transferability of prior knowledge.

sequently, the worker's human capital remains constant and worker knowledge transfers to entrepreneur B, whose productivity z_B increases gradually.

Adoption dynamics. Figure 5b depicts dynamics relating to adoption and poaching. Initially, the worker is matched with entrepreneur A. If productivity z_A and human capital h are sufficiently large, the match finds it optimal to adopt the new technology vintage. In this event, only some of the knowledge in the old technology is transferable to the new vintage. This is captured by a discrete fall in z_A and h upon adoption. In the illustration, the match draws the low transferability state and the respective knowledge in the new vintage are $\underline{\delta} z_A(0)$ and $\underline{\delta} h(0)$. Over time, as the entrepreneur receives exogenous shocks to their productivity z_A and h grow together.[24] When productivity falls below the worker's human capital, learning dynamics imply that the entrepreneur's productivity will catch up to the worker's human capital over time. This gives rise to a volatility harvesting phenomenon, implying that z_A and h grow together. This positive trend growth can be more generally interpreted as incremental follow-up improvements of an existing technology vintage or a reduced form innovation process.

At a random time, entrepreneur B meets with the worker. If it is profitable, they will form a new match and if optimal entrepreneur B will adopt the new technology in the new match. Since entrepreneur B has not used the technology previously, only a random fraction δ of the old knowledge will be transferred to the new vintage.[25] However, since the worker is already experienced in the new technology vintage entrepreneur B can immediately learn from the worker and their productivity level z_B grows even in absence of any productivity shocks. This illustrates the nature of knowledge spillovers in the model that make it relatively cheaper to adopt via external hiring than internally.

Output dynamics. Figure 5c depicts output dynamics under adoption and non-adoption scenarios. The solid path depicts output if the match adopts the new technology vintage while the dashed path depicts output if the match keeps the old technology vintage. Both lines are subject to the same exogenous shocks. Upon adoption output falls initially due to the limited transferability of knowledge between vintages.

[24]In the full model, this process can also be driven by the mean-reversion component of the exogenous stochastic process for productivity if z_A falls below the long-run mean. This captures the notion that learning may be easier the lower the current knowledge level. Conversely, the more productive the entrepreneur the harder it becomes to attain additional knowledge improvements.

[25]In the illustration, the match draws the high transferability state, capturing the assumption that $\mathcal{P}(0,1) > \mathcal{P}(0,0)$ which makes this event more likely when adopting with an experienced worker. This assumption is not strictly necessary, but will be quantitatively important to match the observed poaching patterns in the data.

Since productivity in the new technology vintage is scaled by a factor A, the positive shocks to productivity improve output by a larger amount under the adoption scenario and with time output exceeds the non-adoption counterfactual. This illustrates the dynamics of lower initial output but long-run gains of employing the new technology vintage.

3.8 Calibration

This section discusses key aspects of the calibration. Appendix B and Table B.1 provide further details.

Initial steady-state. Labor market parameters are chosen to be consistent with key labor market statistics from the U.S. In the calibrated initial ergodic steady-state the unemployment rate equals 5%, quarterly employment-to-employment transition rate equals 7.8%, the quarterly job finding rate of unemployed is 72% and the quarterly job separation rate equals 4%. The calibrated productivity process produces a 90-10 differential in TFP of 0.83, which is close to but slightly above the figure stated for manufacturing firms in the U.S. in Syverson (2011).[26] Learning dynamics parameters α and α_u are calibrated to be consistent with Herkenhoff et al. (2024).[27] The discount rate is 8% annually. The value is higher than usual to capture some degree of risk aversion, which is in line with Herkenhoff et al. (2024).

Technology parameters. A crucial factor determining the speed of diffusion of the new technology vintage is the ease at which a non-adopter solo entrepreneur is able to poach a worker from a previous adopter. In the empirical section, I found that flows from previous adopters to newly adopting firms are non-negligible and increase by approximately 15% at the time of adoption. In the model, these poaching flows are modulated by the stochastic process for δ in (4), which is chosen to match our empirical target. Appendix B discusses the strategy.

[26] The 90-10 differential refers to the log difference between the 90th and the 10th percentiles of the TFP distribution.
[27] The steady-state version of my model with $\mathcal{T} = \{0\}$ is closely related to Herkenhoff et al. (2024).

4 Quantitative Model

This section presents the main exercise. While $t < 0$, the economy is at the stationary equilibrium featuring a single vintage, i.e. $\mathcal{T}_t = \{0\}$. At $t = 0$, there is an unanticipated introduction of the new technology vintage and $\mathcal{T}_t = \{0, 1\}$ thereafter. I now study the transition to the new ergodic steady-state.

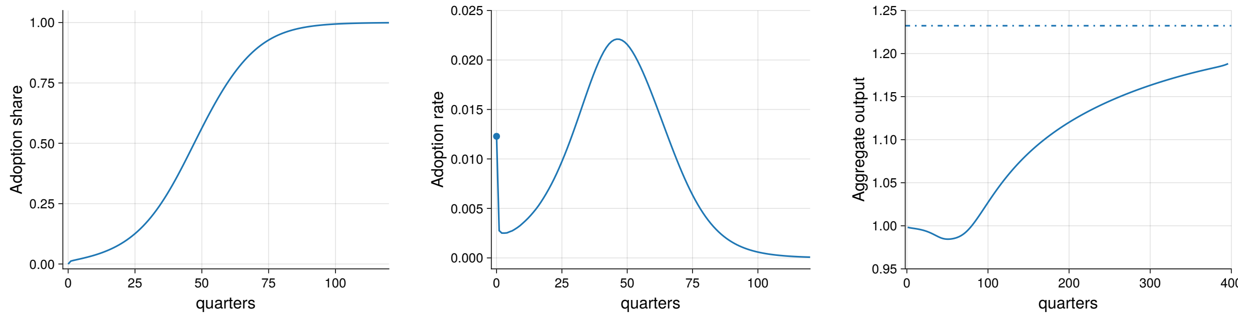
4.1 Diffusion curve

The diffusion of the new, path-breaking technology proceeds in three phases. In the initial phase, approximately 1.5% of entrepreneurs immediately adopt the innovation upon its introduction. The presence of the per-period fixed cost implies that the most productive matches adopt first, in line with empirical evidence (Hampole et al., 2025). During the second phase, adoption rates gradually accelerate over the subsequent decade, reaching 50% penetration within 46 quarters. During the final phase, the remaining, lower productivity firms adopt, with (near-)complete diffusion taking approximately three decades.

Figure 6 illustrates these diffusion dynamics across three dimensions. The left panel depicts a canonical S-shaped adoption curve, demonstrating a slow initial uptake, rapid middle-phase acceleration, and eventual saturation. The center panel presents quarterly adoption rates, showing the underlying changes in technology uptake, with adoption rates peaking after 47 quarters, and the initial jump in adopters after the unexpected arrival of the new technology. The right panel plots the corresponding path for aggregate output over the transition period, showing a J-curve pattern.

The S-shaped diffusion pattern emerges endogenously from the model's random search framework when access to skilled labor is a binding constraint as widely documented in survey data in the case of AI (European Commission, 2020; Rammer & Schubert, 2022). Early in the transition, access to skilled workers and workers with prior experience in the new vintage constitutes a binding constraint to adoption. In the random search framework, contact rates between entrepreneurs and workers are proportional to their existing population, which generates dynamic complementarities similar to epidemiological models of contagion.[28] As a greater population share of entrepreneurs and workers successfully adopts the modern vintage, the likelihood of non-adopters matching with adopters experienced in the new technology

[28] A key difference to models of contagion is that matches have the opportunity to *"infect" themselves*, i.e., adopt internally.



(a) Diffusion curve. **(b)** Adoption rate. **(c)** Aggregate output.

Figure 6: Adoption patterns for entrepreneurs over the transition.

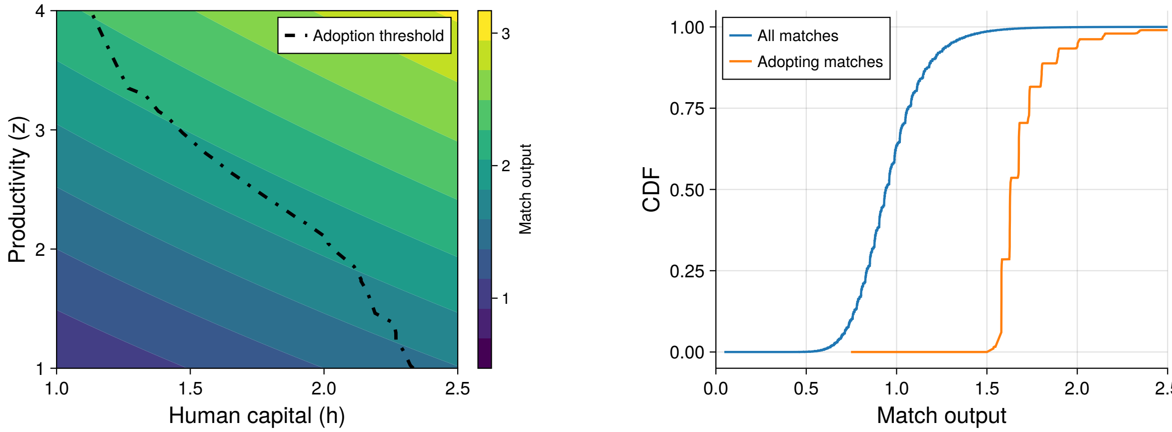
Notes: The left panel plots the share of entrepreneurs that have adopted the new technology after its introduction. The middle panel plots the adoption rate, i.e. the quarterly increase in the number of entrepreneurs. The right panel plots aggregate output over the transition, taking into account changes to knowledge variables, vacancy posting and fixed costs. Time $t = 0$ denotes the time the new technology vintage is introduced.

increases, thereby progressively relaxing the human capital constraint and accelerating diffusion.

Figure 6c exhibits a J-curve pattern, characterized by aggregate output lagging substantially behind technology adoption rates, which has been well documented in the literature on intangible-intensive and general purpose technologies (Brynjolfsson et al., 2021; David, 1990) and aligns with existing theoretical frameworks (Helpman & Trajtenberg, 1994, 1996). Initially, adopting firms have limited expertise in the new technology due to the partial transferability of existing knowledge. Output growth occurs gradually as worker-firm matches mature, and mutual learning generates incremental improvements jointly in the entrepreneur's productivity and the worker's human capital. Dynamics are further modulated by worker mobility: frequent cross-poaching that drives the diffusion disrupts the match-duration dependent knowledge accumulation process. This slows down learning at the individual level. Aggregate output improvements materialize only after approximately 20 years.[29]

Who initiates adoption? Figure 7 illustrates how the diffusion process is initiated. The left panel plots the adoption threshold at impact ($t = 0$)—above which matches adopt the new technology—alongside the

[29] A richer model with e.g. multi-worker firms might feature a shorter lag between output growth and adoption, as pioneering adopters may use their technology advantage to scale up faster and constituting a greater share of output growth. Similarly, allowing for endogenous innovation intensity would have similar effects. Further, note that the model is best viewed as stationarized around a BGP such that aggregate dynamics may deliver small yet positive growth rates.



(a) Adoption threshold and match output.

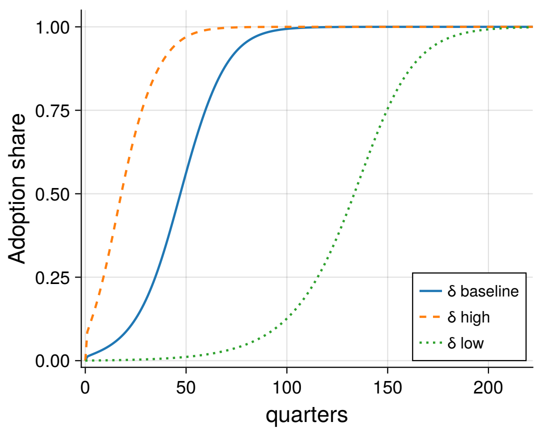
(b) Distributions of initial match output.

Figure 7: Initial adoption impulse at $t = 0$.

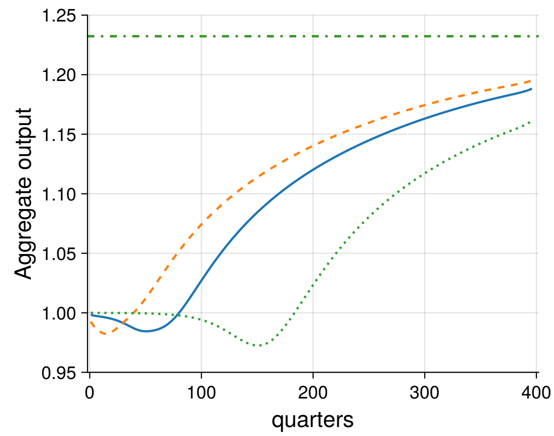
Notes: The left panel shows the contour lines of the initial match output $\pi(z, h \mid 0)$ and the adoption threshold $\underline{z}_t(h)$ at time $t = 0$. Adoption occurs when $z \geq \underline{z}_t(h)$. That is, to the north-east of the dashed line. The right panel plots the conditional CDF of initial match output $\pi(z, h \mid 0)$ for (i) all matches and (ii) all adopting matches. Match output is shown relative to average output.

contour lines for initial output $\pi(z, h \mid 0)$. The figure highlights two main insights. First, more productive matches are first to adopt, consistent with empirical evidence on AI adoption (Hampole et al., 2025). This relationship is further illustrated in the right panel, where the CDF of initial adopters first-order stochastically dominates that of the overall population. This correlation arises because operating the new technology entails a per-period fixed cost and knowledge in the new technology scales in prior expertise.

Second, the adoption threshold has a steeper gradient than the output contour lines, highlighting the role of human capital in shaping the initial adoption decision. Higher human capital directly benefits production and expands the scope for learning, making adoption more attractive despite its initial cost. The intuition is as follows: the more knowledgeable the worker is initially, the more knowledge they build upon after adoption even if this transfer is only partial, as $h_t = \delta h_{t-}$. Consequently, adopting the new technology with a highly skilled worker implies that a match will be more profitable from the start and offers greater opportunities for learning and productivity growth. Further discussion of these competing motives is provided in Appendix A.5.



(a) Diffusion curve.



(b) Aggregate output.

Figure 8: Adoption patterns for different technological parameters for transferability.

Notes: The left panel plots the share of entrepreneurs that adopted the new technology over the transition for three different technological parameters for the transferability of prior knowledge. The solid line represents the calibrated baseline. The dashed line corresponds to a higher transferability technology while the dotted line corresponds to a technology with lower transferability. The right panel plots the respective paths for aggregate output.

Determinants of diffusion speed: Nature of technology. The past has seen a variety of path-breaking technologies with different set of characteristics, ranging from the steam engine during the First Industrial Revolution, electric power generation to ICT, AI, and Electric Vehicles. Through the lens of the framework, a crucial factor determining their speed of diffusion is the degree to which entrepreneurs are constrained by scarce skilled labor. This is closely linked to the transferability of existing skills to new technologies. The more transferable existing skills are to the new technology vintage the less constrained firms are internally and the lower the need for external labor to adopt the new technology, thereby muting the dynamic complementarities. This can be clearly seen in the extreme case in which knowledge is perfectly transferable, i.e. $\overline{\delta} = \underline{\delta} = 1$. In this case, there is no difference between *internal adoption* and *adoption via hiring* for a given z and h, and the role of external workers and labor markets is muted.[30]

To illustrate this dependence, Figure 8a shows adoption patterns under different knowledge transferability regimes, i.e. different values for $\underline{\delta}$ and $\overline{\delta}$. A higher transferability value, i.e. high $\mathbb{E}[\delta]$, now

[30]This is not to say that labor markets will have no role, even in this extreme case. For example, they will be still crucial for enabling more productive firms to scale up and efficiently allocate workers across heterogeneous firms as in Hopenhayn and Rogerson (1993) and Bilal et al. (2022). But these factors are outside the current model.

generates a diffusion curve that may no longer follow an S-shape, with a large fraction of adoption occurring at impact. Firms no longer need to source talent from outside as much and adopt internally. However, adoption is still sluggish as firms require sufficiently high productivity shocks to reach the adoption threshold and solo entrepreneurs require a worker in order to adopt, which requires forming a suitable match in a frictional labor market. Conversely, a lower transferability of existing skills, i.e. small $\mathbb{E}[\delta]$, slows down diffusion as firms are relatively more constrained internally and rely on external hires to relax this constraint. While internal adoption requires that both entrepreneur and worker newly learn operating the technology vintage, adoption via hiring only requires the entrepreneur to newly learn about the technology. Further, the assumption on mutual learning implies that the worker's knowledge is incrementally transferred to the entrepreneur.

These predictions find support in (Adão et al., 2024). Comparing the speed of the diffusion of manufacturing and ICT technologies, the authors show that the ICT revolution required substantive changes in the type of tasks being performed, while the new manufacturing jobs were largely able to reuse existing skills. As a consequence, the speed of diffusion is documented to have been faster for the latter than the former.

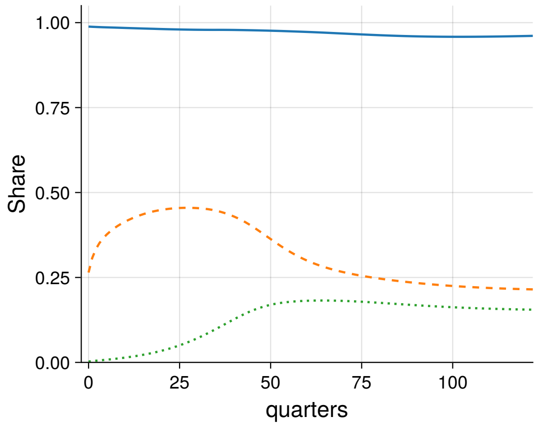
4.2 Importance of worker flows

Access to skilled workers as a binding constraint. A key argument in the preceding sections is that the speed and shape of diffusion critically depends on the extent to which entrepreneurs are constrained in accessing skilled workers experienced in the new technology. This section examines and confirms this channel in the model.

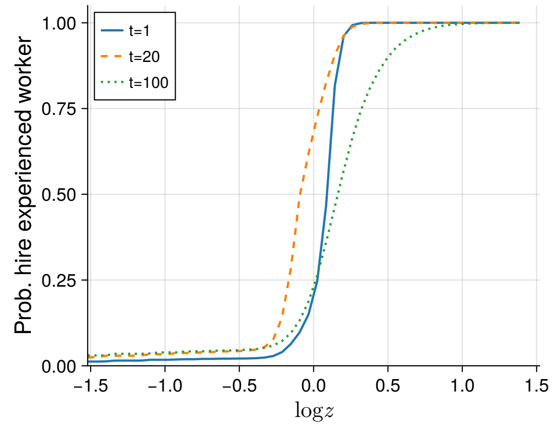
The left panel of Figure 9 shows three different measures for the desire and ability to adopt the new technology over the transition. First, the solid line plots the share of solo, non-adopter entrepreneurs that *would* newly adopt conditional on meeting with a skilled worker. Nearly all entrepreneurs would adopt if they could access an appropriately skilled worker.[31]

Second, the dashed line in the left panel plots the share of solo, non-adopter entrepreneurs that *do*

[31] The moderate fall in the share over the transition captures the compositional effects that the most productive matches will adopt first, implying that the remaining pool of non-adopters is relatively less productive towards the later stages of the transition.



(a) Measures of desire and ability to adopt via poaching.



(b) Ability to poach experienced worker by entrepreneur productivity.

Figure 9: Access to skilled workers as a constraint to technology adoption.

Notes: The left panel plots the fraction of all non-adopter solo entrepreneurs who *would* adopt if matched with an experienced worker (blue, solid), the fraction of non-adopter solo entrepreneurs who *do* poach and adopt *conditional* on meeting an experienced worker (orange, dashed), and the fraction of non-adopter solo entrepreneurs who *do* poach and adopt via poaching an experienced worker. The right panel plots the fraction of non-adopter solo entrepreneurs who are able to hire an experienced worker and adopt the new technology.

newly adopt conditional on meeting with a skilled worker. The difference between the solid and the dashed lines captures the (in-)ability of a solo entrepreneur to successfully poach a worker from a previously adopting firm. The share follows a hump shape over the transition. Initially, only the most productive matches adopt such that pioneering adopters are highly productive, making it hard to poach from. As the technology diffuses to less productive firms poaching becomes relatively easier. At last, as technology-specific knowledge is accumulated within matches the value of continuing a worker-entrepreneur relationship increases and poaching becomes harder again. These dynamics are reflected in the right panel of Figure 9, which shows which entrepreneurs are constrained.[32] Initially, the line shifts to the left indicating that poaching becomes feasible for a larger share of entrepreneurs and then over the long-run shifts back to the right.[33]

Third, the dotted line in the left panel plots the fraction of solo, non-adopter entrepreneurs that

[32] Specifically, lower values imply a larger degree of being constrained.

[33] The lines shift upwards over the transition as matches with the new technology vintage separate over time and unemployment represents slower reallocation dynamics that are accessible even to low productivity entrepreneurs.

newly adopt with a skilled worker, unconditionally. The difference between the dashed and the dotted line captures the probability of meeting with a skilled, experienced worker. The increase in the dotted line thus captures the rising probability of encountering an experienced worker over the transition. This captures the fact that the constraint on skilled labor relaxes over the transition.

Which worker flows matter? This section examines how the economy overcomes the constraint on experienced labor and specifically which type of worker flows matter at different points in the transition. To this end, I decompose the relative importance of the distinct adoption channels. Following Definition 3, I categorize adoption by entrepreneurs into two channels: (i) *internal adoption* (ii) *adoption via hiring*. The latter channel is further distinguished by whether a solo entrepreneur forms the new match with a poaching hire or with an unemployed worker.[34]

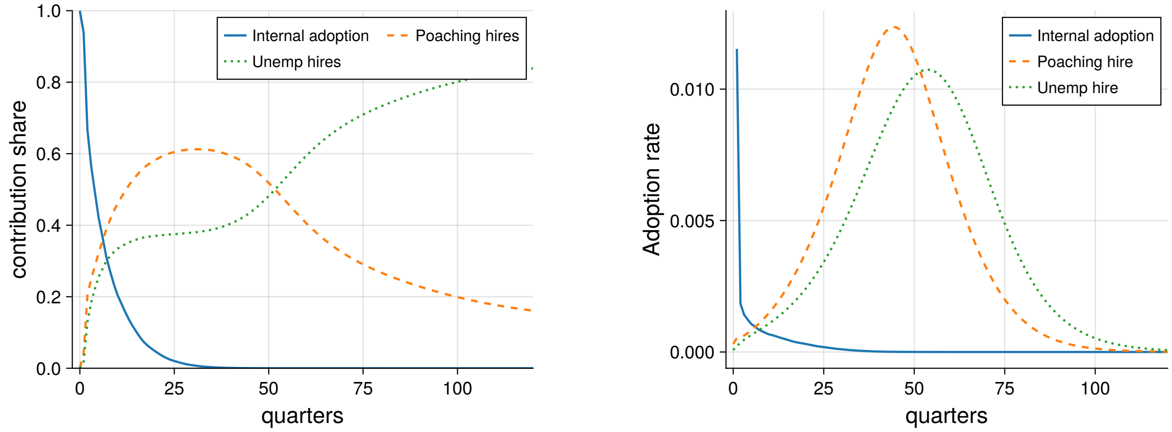
Figure 10 illustrates the relative contributions of those channels over the first three decades of the transition period. The left panel plots the share of newly adopting entrepreneurs coming from each of the three channels. The right panel plots the absolute quarterly rate at which entrepreneurs newly adopt. The sum of the three lines equals the adoption rate plotted in Figure 6b.[35] The importance of each channel varies over the transition.

Internal adoption. Initially, internal adoption is the primary driver behind the diffusion. However, the importance of this channel quickly diminishes as free-riding incentives begin to dominate. That is, as the share of experienced workers increases in the population the option value to wait to adopt via hiring such a worker becomes increasingly more attractive (see Appendix D.1). This, reduces the willingness of entrepreneurs to choose the more costly internal adoption.

Poaching hires. During the accelerating phase of diffusion, adoption is primarily driven by poaching. The shape of this curve mirrors Figure 9a, reflecting the time-varying ease of hiring workers away from other firms.

[34]Likewise, adoption by workers is categorized into two channels: (i) internal adoption and (ii) adoption via training following Definition 4. The latter channel is again distinguished by a poaching or unemployment hire.

[35]The notion of a decomposition is subtle. Conceptually, all adopters can be traced back to internal adopters who are necessary to set the process in motion. That is, all contagion is traced back to a set of *patient zeros*. Contribution shares as used here refer to the breakdown of the various channels of adoption. For example, it does not take explicitly into account that adoption via poaching will by itself trigger more adoption via unemployment hires in the future.



(a) Contribution shares by adoption channels.

(b) Adoption rates by adoption channels.

Figure 10: Contributions by adoption channels (entrepreneur).

Notes: The left panel plots the share of newly adopting entrepreneurs attributable to each of three channels: internal adoption (blue, solid), poaching hires (orange, dashed), and unemployment hires (green, dotted). The right panel plots the total quarterly rate of newly adopting entrepreneurs attributable to each respective channel. Further details are provided in Appendix A.6.

Unemployment hires. In the later stages, unemployment hires become the dominant channel. This process captures the slow-moving reallocation forces that take time to materialize. As poaching becomes increasingly difficult toward the end, unemployment hires emerges as the primary source of diffusion. In the model, reallocation via unemployment captures in a reduced form general slow-moving reallocation channels (see Appendix D.2).

4.3 Divergence between U.S. and EU economies

Growth rates in the EU and U.S. have diverged substantially over the past two decades, by approximately one percentage point per year. Draghi (2024) provides a comprehensive investigation of the sources of this divergence and emphasizes one primary factor that account for most of the trend: differential productivity growth in information and communication technologies (ICT). While the EU, and Germany in particular, maintain a productivity lead in mature industries characterized by incremental innovation, European economies lag behind in technology-intensive sectors exhibiting rapid technical change. This pattern is notable given the well-documented differences in labor market institutions between these economies

Table 1: Initial steady-state properties in high mobility (U.S.) and low mobility & long tenure (EU) calibration.

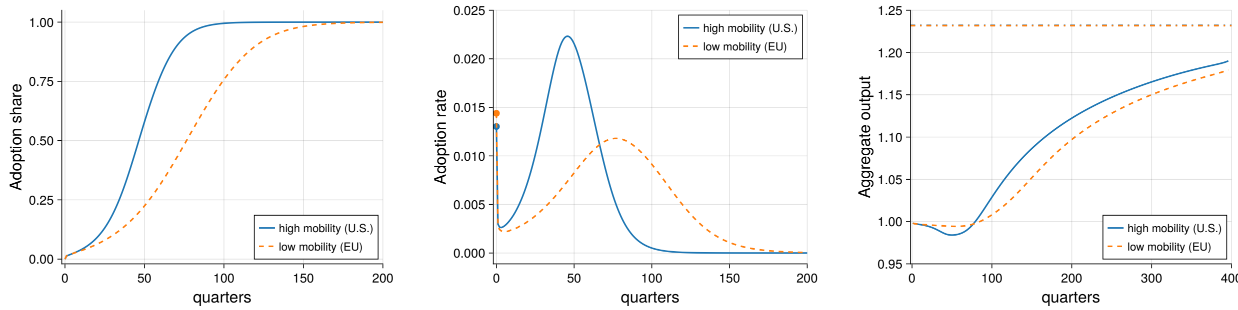
	E-U rate	E-E rate	U-E rate	Initial s.s.	Final s.s.	Growth
High mobility (U.S.)	0.04	0.078	0.72	100%	123%	23.4%
Low mobility (EU)	0.02	0.04	0.38	102%	126%	23.4%

Notes: Initial steady-state, i.e. before the introduction of the new technology, is relative to the high mobility (U.S.) economy. The growth rate is computed between initial and final steady-state. E-U denotes employment-to-unemployment rate, E-E the employment-to-employment rate, and U-E the unemployment-to-employment rate. All rates are quarterly.

(Borowczyk-Martins, 2025).

With my model establishing worker mobility as crucial for technology diffusion, I examine whether labor market institutions have contributed to the EU-U.S. divergence and can help explain why Europe excels in incremental innovation but lags in rapidly-evolving technological sectors. I focus on two defining features of European labor markets: lower worker mobility and longer tenure jobs (Borowczyk-Martins, 2025). Specifically, I re-calibrate the quantitative model to match the reduced separation rates and matching efficiency observed in many European economies, while holding the unemployment rate constant, thereby capturing cross-country differences in labor market institutions in reduced form. This calibration yields lower employment-to-employment (E-E) and unemployment-to-employment (U-E) transition rates, alongside the longer job tenures characteristic of European labor markets.

Steady-state. I begin by comparing the old steady-state, i.e. $t < 0$, of the low mobility, long tenure (EU) calibration with the high mobility (U.S.) baseline. Results are shown in Table 1. Lower mobility rates and longer job duration are associated with modestly higher steady-state output, while total output growth from the new technology is approximately equal in both calibrations. Intuitively, longer job duration from reduced separation rates facilitates on-the-job learning through the mutual transfer of knowledge. That is, the learning process defined in (3) benefits from match duration. These steady-state advantages align with Europe's documented success in industries requiring incremental innovation, where stable employment relationships facilitate gradual productivity improvements.



(a) Diffusion curve. **(b)** Adoption rate. **(c)** Aggregate output.

Figure 11: Adoption patterns for high mobility (U.S.) and low mobility & long tenure (EU) calibrations.

Notes: High-mobility calibration corresponds to the U.S. calibration (blue, solid). Low-mobility & long tenure job calibration corresponds to the EU calibration (orange, dashed). The left panel plots the share of entrepreneurs that have adopted the new technology after its introduction. The middle panel plots the quarterly adoption rate. The right panel plots aggregate output over the transition, taking into account changes to knowledge variables, vacancy posting and fixed costs. Time $t = 0$ denotes the time the new technology vintage is introduced. Technological parameters are constant across countries. Only labor market parameters are re-calibrated.

Transition dynamics. The transition dynamics reveal a striking contrast with the steady-state results. Figure 11 displays adoption and output paths for both calibrations, showing that diffusion in the low mobility, long tenure (EU) calibration proceeds two-thirds more slowly than in the U.S. baseline. These more sluggish dynamics reflect fundamental differences in how labor markets allocate scarce technological expertise. Solo entrepreneurs in the EU calibration face binding constraints on accessing workers experienced with the new technology: instead of poaching, they must rely on the slower process of exogenous separations to attract skilled labor. Moreover, longer job tenures impede even this reallocation channel, as workers cycle through unemployment less frequently. These frictions compound to increase the technology's half-life by approximately 65% or $7-8$ years relative to the U.S. baseline.

The output implications of differential diffusion patterns are shown in Figure 11c. At first, U.S. output falls slightly relative to the EU calibration, as firms in the U.S. are faster to adopt the new technology, yet lack the experience initially. As productivity gains in the new technology materialize, an output gap between the U.S. and EU, $Y_t^{US} - Y_t^{EU}$, emerges. Once established, the output gap exhibits strong persistence and convergence remains minimal absent institutional change. Aggregate output can thus be an inadequate contemporaneous measure for the future economic performance early in the transition. In

Table 2: Net Present Value of Arrival of New Vintage

	high mobility (U.S.)	low mobility (EU)
Change in Welfare (% GDP)	26%	17.5%
Half life (quarters)	46.4	76.5

contrast, adoption rates are a more informative leading indicator.

Welfare. How do the differential adoption rates and output paths under the respective labor market calibrations map into welfare? To answer this question, I next compute the change in welfare attributable to the arrival of the new technology vintage for each calibration, respectively. Since utility is transferable, the changes in welfare correspond to changes in net output relative to initial GDP.[36] That is,

$$\Delta W_0 = \int_0^{+\infty} e^{-rt} \left(\frac{Y(t) - Y(0)}{Y(0)} \right) dt. \tag{9}$$

Numerical results are summarized in the Table 2. The introduction of a new technology vintage increases the net present value of welfare in terms of initial steady-state GDP by 26% in the high mobility (U.S.) economy and only by 17.5%, one-third lower, in the lower mobility (EU) economy. These effects are primarily driven by a longer half-life of diffusion: reaching 50% penetration takes about 7 to 8 years (two-thirds) longer in the low mobility, long tenure economy.[37] These results suggest that labor market institutions can have sizable effects on adoption speeds and welfare, thereby proposing a quantitative mechanism for the diverging transatlantic productivity gaps.

[36]Net output subtracts vacancy posting costs and the per-period adjustment fixed cost when using the new technology vintage.

[37]The conclusions of this exercise draw parallels to Ljungqvist and Sargent (1998), who similarly find small differences in steady-state but very different (unemployment) dynamics between EU and U.S. economies. While their focus was on the welfare system, the present paper shares the point that the steady-state might conceal welfare losses from limited agility.

5 Policy and Efficiency

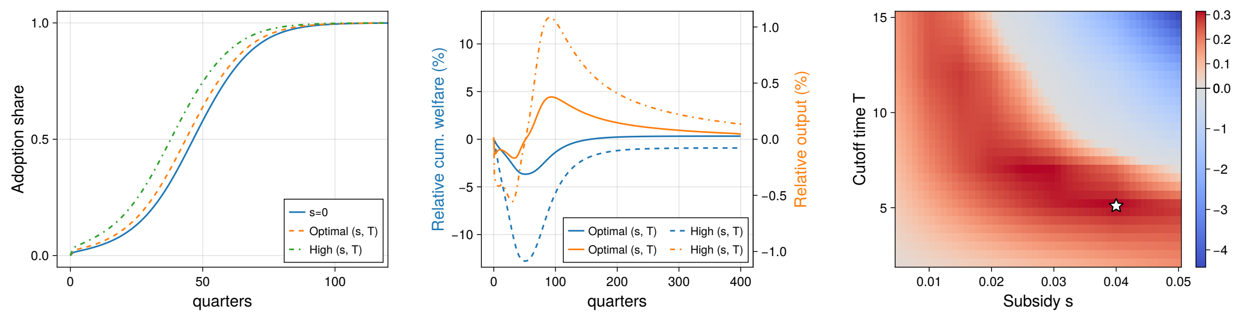
5.1 Inefficiencies

As a corollary of the sequential auction bargaining protocol all decisions in the model are bilaterally efficient, meaning inefficiencies take the form of spillover externalities to third parties. First, matches do not internalize that the mobility of the worker creates value to other entrepreneurs through subsequent job-to-job transitions. For example, when deciding whether to adopt a new technology vintage, matches do not consider that the mobility of their worker can facilitate adoption by solo entrepreneurs in the future, possibly leading to inefficiently low internal adoption. Second, the search and matching framework gives rise to standard congestion externalities as in Hosios (1990). That is, entrepreneurs do not internalize that their vacancy posting increases the rate at which a single worker is contacted $p(\theta_t)$ and decreases the rate at which a single solo entrepreneur meets a worker $q(\theta_t)$. Likewise, workers do not internalize that their participation in search increases $q(\theta_t)$ while lowering the job $p(\theta_t)$. This may lead to inefficiently high vacancy posting. Appendix G illustrates these spillover externalities in an analytical model featuring three firms and a single worker.

5.2 Adoption subsidy

Before discussing the fully optimal policy, this section considers a particular adoption subsidy designed to address the above inefficiencies: a two-part policy (s, T). In this scheme, output from using the new technology vintage is subsidized at a rate s, but the scheme expires after T periods. The subsidy is financed by uniform revenue taxes and the budget is balanced in every period. That is, the after-tax flow revenue of an entrepreneur-worker match is

$$\mathcal{R}(z, h \mid \tau) = \begin{cases} (1 + s - tax_t)\pi(z, h \mid \tau) & \text{if } 0 \leq t \leq T \text{ and } \tau = 1, \\ (1 - tax_t)\pi(z, h \mid \tau) & \text{if } 0 \leq t \leq T \text{ and } \tau = 0, \\ \pi(z, h \mid \tau) & \text{if } t > T. \end{cases}$$



(a) Diffusion curve. **(b)** Output and cumulative welfare. **(c)** Differential welfare ΔW_0-ΔW_0^*(%).

Figure 12: Adoption subsidy with expiration time.

Notes: The left panel plots the adoption shares for three policies: (i) no policy baseline, i.e. $s = 0$ (ii) the optimal (s, T) menu and (iii) a high $(s, T) = (0.04, 12)$ schedule delivering faster adoption. The middle panel shows relative cumulative welfare (left axis) and relative output (right axis) over the transition. The right panel shows the differential net present value ΔW_0-ΔW_0^* for different policy menus (s, T). The welfare gain ΔW_0 as defined in (9) is normalized by the welfare gain without the subsidy, ΔW_0^*. The white star denotes the optimal combination of (s^*, T^*).

This subsidy generates incentives to adopt the new technology by increasing the expected revenues after adoption. Specifically, the subsidy targets early adoption, as the effective subsidy declines over the transition and expires at some $t = T$. Balance budget implies that in every period $0 \leq t \leq T$

$$tax_t = s \cdot \frac{\int \pi(z, h \mid 1) dG_t^e(z, h \mid 1) + \int \pi(z, 0 \mid 1) dG_t^v(z \mid 1)}{\sum_\tau \int \pi(z, h \mid \tau) dG_t^e(z, h \mid \tau) + \sum_\tau \int \pi(z, 0 \mid \tau) dG_t^v(z \mid \tau)}.$$

Optimal subsidy & expiration-time policy. Figure 12 illustrates how different policy schemes (s, T) shape adoption patterns, aggregate output, and cumulative welfare. The left panel demonstrates that positive subsidies accelerate technology diffusion by incentivizing early adoption. The middle panel highlights a critical intertemporal trade-off: accelerated adoption generates short-run output losses (orange, right axis) but yields long-run output gains, resulting in welfare losses over the initial quarters relative to the no-subsidy baseline (blue, left axis). The magnitude of this trade-off depends on diffusion speed—overly rapid diffusion, driven by large subsidies, produces prohibitive short-run output losses that future gains cannot offset, while moderate acceleration minimizes initial losses and generates positive, though small, long-run welfare gains (solid, blue).

The right panel maps the welfare differential net of the no-subsidy baseline from technology introduction, ΔW_0-ΔW_0^*, across the (s, T) space, with the optimal combination marked by the white star. Welfare exhibits non-monotonicity in both dimensions. While higher subsidies, s, and longer durations, T, both increase the net present value of adopting the new technology, thereby accelerating diffusion, this non-monotonicity reveals that transitions can be *too* fast. The optimal policy features moderate subsidies targeted at early adopters with relatively short expiration horizons.

Transitions can be too fast. These results highlight an important policy insight: accelerating diffusion is not unconditionally welfare-improving. Intuitively, as subsidies expand, they induce adoption by progressively less productive matches. Since productivity in the new technology scales with existing capabilities, these marginal adopters lack sufficient profitability to justify the switch to the new technology and incur the per-period fixed cost. This mechanism contributes to the initial decline in output when subsidies are either too large or maintained for too long. From a welfare perspective, these firms would benefit from delayed adoption, allowing them to first absorb knowledge spillovers from workers gaining experience at high-productivity early adopters. This learning externality creates a natural motive for measured, rather than maximal, diffusion speed. Put differently, the subsidy leads to excessive *internal* adoption instead of the cheaper adoption through hiring.

Relation to inefficiencies. The subsidy addresses the two main externalities defined in the preceding discussion. First, matches fail to internalize the positive knowledge spillovers their workers generate when moving to future adopters when they decide whether to upgrade to the new technology. Second, firms do not fully internalize the negative congestion effects their labor market participation imposes. The strength of these externalities varies over the transition: knowledge spillovers are strongest early in the transition when the pool of non-adopters is largest, while congestion effects are strongest when competition for and poaching of diffusive workers intensifies. A time-limited subsidy targeted at pioneering adopters balances these intertemporal trade-offs.

5.3 Optimal policy

Having discussed a particular adoption subsidy, I now compute the set of optimal taxes and subsidies which implement the social planner allocation.[38] Implementation requires two types of instruments, corresponding to the two margins of decision-making. For discrete choices—all decisions except vacancy posting—lump-sum, match-specific transfers equate private and social valuations of matches, solo entrepreneurs, and the unemployed. When these valuations align, the indicator policy functions in the competitive equilibrium coincide with those of the social planner. For vacancy posting, the sole continuous choice, implementation additionally requires a marginal tax to align first-order conditions. In summary, discrete choices require equalizing valuations through transfers while continuous choices require equalizing both valuations and marginal conditions.

Proposition 1. *The decentralized equilibrium is generically inefficient. Let tilde variables represent the social planner counterparts to the competitive equilibrium variables. The optimal revenue-neutral transfers and tax are*

$$T_t^e(z, h \mid \tau) = \Gamma_t^e(z, h \mid \tau) - \frac{\xi}{\tilde{\mathbb{X}}_t}\mu_t$$
$$T_t^u(h \mid \tau) = \Gamma_t^u(h \mid \tau) - \frac{1}{\tilde{\mathbb{X}}_t}\mu_t,$$

while the optimal marginal tax on vacancy posting is

$$\varsigma_t^v = \frac{1}{\tilde{\mathbb{V}}_t}\omega_t,$$

where the positive spillover externalities are defined as

$$\Gamma_t^e(z, h \mid \tau) = \xi p(\tilde{\theta}_t) \sum_{\tau' \in \mathcal{T}_t} \int \tilde{\Delta}_t^{EE}(z, h, z' \mid \tau, \tau')^{+} \frac{\tilde{v}_t(z' \mid \tau')}{\tilde{\mathbb{V}}_t} d\tilde{G}_t^v(z' \mid \tau')$$
$$\Gamma_t^u(h \mid \tau) = p(\tilde{\theta}_t) \sum_{\tau' \in \mathcal{T}_t} \int \tilde{\Delta}_t^{UE}(h, z' \mid \tau, \tau')^{+} \frac{\tilde{v}_t(z' \mid \tau')}{\tilde{\mathbb{V}}_t} d\tilde{G}_t^v(z' \mid \tau'),$$

and the congestion externality wedges are, letting $\eta(\theta) = p'(\theta)\theta/p(\theta)$ denote the elasticity of the aggregate

[38] This generalizes Fukui and Mukoyama (2025) to a setting with heterogeneous human capital, learning and technology choice.

matching function with respect to the aggregate search input $\mathbb{X}t$

$$\mu_t = \eta(\tilde{\theta}_t)\left(\sum_{\tau\in\mathcal{T}_t}\int \Gamma^e_t(z,h\mid\tau)d\tilde{G}^e_t(z,h\mid\tau) + \sum_{\tau\in\mathcal{T}_t}\int \Gamma^u_t(h\mid\tau)d\tilde{G}^u_t(h\mid\tau)\right)$$

$$\omega_t = \big(1-\eta(\tilde{\theta}_t)\big)\left(\sum_{\tau\in\mathcal{T}_t}\int \Gamma^e_t(z,h\mid\tau)d\tilde{G}^e_t(z,h\mid\tau) + \sum_{\tau\in\mathcal{T}_t}\int \Gamma^u_t(h\mid\tau)d\tilde{G}^u_t(h\mid\tau)\right).$$

Vacancy posting. Proposition 1 shows that vacancy posting is inefficiently high in the competitive equilibrium and should be taxed at a strictly positive rate $\varsigma^v_t > 0$.[39] Solo entrepreneurs appropriate the entire surplus gain from poaching and hiring, while imposing congestion externalities on other labor market participants. To correct this inefficiency, the social planner sets the marginal tax equal to the per-vacancy congestion externality $\varsigma^v_t = \omega_t/\tilde{\mathbb{V}}_t$. Since all solo entrepreneurs generate identical congestion externalities per vacancy, this tax *rate* is uniform.

Congestion externality. Employed and unemployed workers exert negative congestion externalities, scaled by their contribution to the aggregate search input. As workers only make discrete choices simple lump-sum taxes, $\frac{\xi}{\tilde{\mathbb{X}}_t}\mu_t$ and $\frac{1}{\tilde{\mathbb{X}}_t}\mu_t$ respectively, corrects for this.[40]

Positive spillovers. Matched entrepreneurs exert positive spillover effects Γ^e_t via job-to-job mobility, which are not internalized in the present bargaining protocol and should be subsidized.[41] These spillover effects are increasing in labor market tightness, θ_t, and depend on the expected surplus gains from hiring. Within a technology vintage, the spillover gains favor the formation of bridge jobs that help diffuse knowledge and reallocate to higher value jobs. For example, the social planner values more strongly the creation of high learning gain jobs as the human capital acquired on the job can later be used to be diffused to other entrepreneurs.[42] Across technology vintages, the spillover gains favor the adoption of the new technology if the worker is likely to facilitate the adoption by other entrepreneurs through job-to-job movements.

[39] The total cost of posting v vacancies is time-dependent and equals $c(v) + \varsigma^v_t v$.

[40] If workers chose their search effort, the correction would be similar to the marginal vacancy posting tax on entrepreneurs.

[41] This is in part an implication of the particular bargaining protocol. The generalization of the sequential auction bargaining protocol by Cahuc et al. (2006) allows matches to, at least partially, internalize these knowledge spillovers. Likewise, buyout offers are other institutional designs that allow some internalization of these externalities. However, these features would introduce other inefficiencies as solo entrepreneurs would be subject to the well-known hold-up problem.

[42] A clean example is shown in the analytical model in G.

Employer-to-employer transitions. The movement of a worker from a match (z, h, τ) to a solo entrepreneur (z', τ') is shaped by the optimal transfers through two channels. First, the optimal transfers direct workers toward firms with larger spillover externalities—typically those enabling technology adoption or worker-firm learning.[43] Second, the marginal tax ς_t^v directs moves to entrepreneurs with high vacancy-posting when solo, since these entrepreneurs have a lower option value of being solo in the social planner allocation as taxes are proportional to overall recruitment effort.

Technology adoption. The adoption decision depends on the differential technology-specific lump-sum taxes when operating the old versus the new technology. These transfers capturing the positive spillover externalities are endogenously determined by worker mobility patterns. For example, if workers are relatively more likely to move to other employers after adoption, then the optimal transfers favor the internal adoption of the new technology by inducing matches to internalize the knowledge spillover.

Unemployment-to-employment transition / separation. The movement between employment and unemployment is shaped by the optimal transfers through two channels. On the one hand, the transfers capturing spillover externalities, conditional on the state (h, τ), tend to be relatively larger when unemployed, as it is cheaper for entrepreneurs to hire out of unemployment than to poach a matched worker. This favors unemployment, i.e. lower U-E and higher E-U rates. On the other hand, the lump-sum taxes targeted to correct for congestion externalities are larger when unemployed if $\xi < 1$, the empirically relevant case, which favors employment.

Role of learning. The presence of learning and technology diffusion breaks the tight connection between productivity and marginal value ladders as commonly used in job ladder models (Fukui & Mukoyama, 2025). As a result, it no longer necessarily holds that high productivity matches are overvalued in the competitive equilibrium relative to the social planner, as long as workers have the potential to diffuse knowledge elsewhere. In absence of knowledge spillovers, high productivity firms are located at the top of the value ladder. As workers at those firms are in the most valuable match, they do not move elsewhere and therefore spillover effects as small. In the presence of diffusive labor mobility, it could even be optimal to subsidize these high revenue firms, especially those experienced in the new technology vintage.

[43] However, note that this is a complex object. What matters is the *relative* likelihood of moving to high-spillover matches which itself depends on the patterns of employer-to-employer transitions.

6 Concluding remarks

This paper establishes worker mobility as a central determinant of how path-breaking technologies diffuse through the economy, particularly during the early stages of the technology lifecycle and when existing knowledge is only partially transferable to the newer technology. I show how skilled labor scarcity can create a binding constraint on technology adoption that endogenously generates the empirically-relevant S-shaped diffusion curve and J-curve productivity patterns. Early in the transition, access to workers with new-vintage experience constrains adoption. As diffusion progresses and the pool of experienced workers expands, this constraint relaxes and adoption accelerates, progressively increasing the likelihood that non-adopters match with adopters. Accelerating dynamics are primarily driven by increased poaching. The analysis highlights a fundamental trade-off facing policymakers: low mobility and long tenure labor markets facilitate incremental learning in steady-state but impede the technological agility required during transformative transitions, potentially contributing to the divergence in U.S. and EU productivities in technology-intensive sectors. Subsidies aimed at incentivizing adoption are optimally targeted at early adopters to initiate the diffusion process and strike an intertemporal balance between competing externalities.

The framework developed in the present paper opens several avenues for future research. First, the current version assumes the exogenous arrival of a single new vintage. This could be relaxed along two dimensions in the pursuit of a unified account of growth: endogenizing the sequential development of radical technologies and the co-existence of multiple, possibly competing vintages. Second, the model could be extended to multi-worker firms to study the interaction between internal labor markets and organizational scale, while also capturing within-firm knowledge spillovers that may create heterogeneous technology intensity within firms. Third, the framework could be enriched to evaluate specific policy levers such as university education, dynamic entry and creative destruction as well as concrete training policies and international migration and visa programs. Finally, an important open question is the role of the bargaining protocol and surplus sharing between agents. Relaxing the assumption on bilateral efficiency could be an interesting avenue.

References

Abis, **S.**, & **Veldkamp**, **L.** (2023). The changing economics of knowledge production. *The Review of Financial Studies.*

Acemoglu, **D.**, **Akcigit**, **U.**, & **Celik**, **M. A.** (2022a). Radical and incremental innovation: The roles of firms, managers, and innovators. *American Economic Journal: Macroeconomics.*

Acemoglu, **D.**, **Autor**, **D.**, **Hazell**, **J.**, & **Restrepo**, **P.** (2022b). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics.*

Adão, **R.**, **Beraja**, **M.**, & **Pandalai-Nayar**, **N.** (2024). Fast and slow technological transitions. *Journal of Political Economy Macroeconomics.*

Adenbaum, **J.**, **Babalievsky**, **F.**, & **Jungerman**, **W.** (2024). *Learning on the job.*

Aghion, **P.**, **Howitt**, **P.**, & **Violante**, **G. L.** (2002). General purpose technology and wage inequality. *Journal of Economic Growth.*

Akcigit, **U.**, **Caicedo**, **S.**, **Miguelez**, **E.**, **Stantcheva**, **S.**, & **Sterzi**, **V.** (2018). *Dancing with the stars: Innovation through interactions* (NBER Working Paper).

Atkeson, **A.**, & **Kehoe**, **P. J.** (2007). Modeling the transition to a new economy: Lessons from two technological revolutions. *American Economic Review.*

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

Benhabib, **J.**, **Perla**, **J.**, & **Tonetti**, **C.** (2021). Reconciling models of diffusion and innovation: A theory of the productivity distribution and technology frontier. *Econometrica*, *89*(5), 2261–2301.

Bilal, **A.**, **Engbom**, **N.**, **Mongey**, **S.**, & **Violante**, **G. L.** (2022). Firm and worker dynamics in a frictional labor market. *Econometrica*, *90*(4).

Borowczyk-Martins, **D.** (2025). *Employer-to-employer mobility and wages in europe and the united states* (IZA Discussion Papers).

Bresnahan, **T. F.**, & **Trajtenberg**, **M.** (1995). General purpose technologies 'engines of growth'? *Journal of Econometrics.*

Brynjolfsson, **E.**, **Chandar**, **B.**, & **Chen**, **R.** (2025). *Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence* (tech. rep.).

Brynjolfsson, **E.**, **Rock**, **D.**, & **Syverson**, **C.** (2021). The productivity j-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics.*

Buera, **F. J.**, & **Oberfield**, **E.** (2020). The global diffusion of ideas. *Econometrica.*

Cahuc, **P.**, **Postel-Vinay**, **F.**, & **Robin**, **J.-M.** (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica.*

Caselli, **F.** (1999). Technological revolutions. *American Economic Review*.

Cestone, **G.**, **Fumagalli**, **C.**, **Kramarz**, **F.**, & **Pica**, **G.** (2023). Exploiting growth opportunities: The role of internal labour markets. *The Review of Economic Studies*, *91*(5), 2676–2716.

Chari, **V. V.**, & **Hopenhayn**, **H.** (1991). Vintage human capital, growth, and the diffusion of new technology. *Journal of Political Economy*, *99*(6).

David, **P. A.** (1990). The dynamo and the computer: An historical perspective on the modern productivity paradox. *The American Economic Review*.

Davis, **S. J.**, & **Haltiwanger**, **J.** (2014). *Labor market fluidity and economic performance* (tech. rep.).

De Ridder, **M.** (2024). Market power and innovation in the intangible economy. *American Economic Review*.

Decker, **R. A.**, **Haltiwanger**, **J.**, **Jarmin**, **R. S.**, & **Miranda**, **J.** (2016). Declining business dynamism: What we know and the way forward. *American Economic Review*.

Draghi, **M.** (2024). The future of european competitiveness: A competitiveness strategy for europe.

Engbom, **N.** (2023). Misallocative growth.

European Commission. (2020). *European enterprise survey on the use of technologies based on artificial intelligence – final report*. Publications Office.

Fukui, **M.**, & **Mukoyama**, **T.** (2025, June). *Efficiency in job-ladder models* (Working Paper No. 33910). National Bureau of Economic Research.

Hampole, **M.**, **Papanikolaou**, **D.**, **Schmidt**, **L. D.**, & **Seegmiller**, **B.** (2025). *Artificial intelligence and the labor market* (Working Paper). National Bureau of Economic Research.

Helpman, **E.**, & **Trajtenberg**, **M.** (1994). *A time to sow and a time to reap: Growth based on general purpose technologies* (NBER Working Papers). National Bureau of Economic Research, Inc.

Helpman, **E.**, & **Trajtenberg**, **M.** (1996). *Diffusion of general purpose technologies* (Working Paper). National Bureau of Economic Research.

Herkenhoff, **K.**, **Lise**, **J.**, **Menzio**, **G.**, & **Phillips**, **G. M.** (2024). Production and learning in teams. *Econometrica*.

Hombert, **J.**, & **Matray**, **A.** (2024). Innovation booms, easy financing, and human capital accumulation. *NBER Working Paper*.

Hopenhayn, **H.**, & **Rogerson**, **R.** (1993). Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*.

Hosios, **A. J.** (1990). On the efficiency of matching and related models of search and unemployment. *The Review of Economic Studies*, *57*(2), 279–298.

Humlum, **A.** (2019). Robot adoption and labor market dynamics.

Jaravel, **X.**, **Petkova**, **N.**, & **Bell**, **A.** (2018). Team-specific capital and innovation. *American Economic Review*.

Jarosch, **G.**, **Oberfield**, **E.**, & **Rossi-Hansberg**, **E.** (2021). Learning from coworkers. *Econometrica*, *89*(2), 647–676.

Jingnan, **L.** (2023). *Worker mobility, knowledge diffusion, and non-compete contracts.*

Jones, **B. F.** (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*.

Jovanovic, **B.**, & **Nyarko**, **Y.** (1996). Learning by doing and the choice of technology. *Econometrica*, *64*(6), 1299–1310.

Jovanovic, **B.**, & **Rousseau**, **P. L.** (2005). Chapter 18 - general purpose technologies. In **P. Aghion** & **S. N. Durlauf** (Eds.).

Kalyani, **A.**, **Bloom**, **N.**, **Carvalho**, **M.**, **Hassan**, **T.**, **Lerner**, **J.**, & **Tahoun**, **A.** (2025). The diffusion of new technologies*. *The Quarterly Journal of Economics*.

Kogan, **L.**, **Papanikolaou**, **D.**, **Schmidt**, **L. D. W.**, & **Seegmiller**, **B.** (2021). Technology, vintage-specific human capital, and labor displacement: Evidence from linking patents with occupations. *NBER Working Paper*.

Koike-Mori, **Y.**, **Maruyama**, **T.**, & **Okumura**, **K.** (2024). *Inventor mobility, knowledge diffusion, and growth.*

Lichtinger, **G.**, & **Hosseini Maasoum**, **S. M.** (2025). Generative ai as seniority-biased technological change: Evidence from u.s. résumé and job posting data. *Available at SSRN 5425555*.

Lise, **J.**, & **Robin**, **J.-M.** (2017). The macrodynamics of sorting between workers and firms. *American Economic Review*, *107*(4).

Ljungqvist, **L.**, & **Sargent**, **T. J.** (1998). The european unemployment dilemma. *Journal of Political Economy*.

Lucas, **R. E.**, & **Moll**, **B.** (2014). Knowledge growth and the allocation of time. *Journal of Political Economy*, *122*(1), 1–51.

Luttmer, **E. G. J.** (2007). Selection, growth, and the size distribution of firms*. *The Quarterly Journal of Economics*.

Mokyr, **J.** (2002). *The gifts of athena: Historical origins of the knowledge economy*. Princeton University Press.

Nelson, **R. R.**, & **Phelps**, **E. S.** (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*.

Perla, **J.**, & **Tonetti**, **C.** (2014). Equilibrium imitation and growth. *Journal of Political Economy*.

Perla, **J.**, **Tonetti**, **C.**, & **Waugh**, **M. E.** (2021). Equilibrium technology diffusion, trade, and growth. *American Economic Review*, *111*(1).

Postel-Vinay, **F.**, & **Robin**, **J.-M.** (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, *70*(6), 2295–2350.

Rammer, **C.**, & **Schubert**, **T.** (2022). *Dokumentation zur Innovationserhebung 2017 bis 2021* (ZEW-Dokumentation). ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung.

Sampson, **T.** (2015). Dynamic selection: An idea flows theory of entry, trade, and growth *. *The Quarterly Journal of Economics*.

Schoefer, **B.** (2025, June). *Eurosclerosis at 40: Labor market institutions, dynamism, and european competitiveness* (Working Paper No. 33975). National Bureau of Economic Research.

Schubert, **G.**, **Stansbury**, **A.**, **Taska**, **B.**, & **Dey**, **M.** (2025). *Employer concentration and wages: Evidence from large firms' hiring shocks* (tech. rep.).

Schumpeter, **J. A.** (1939). *Business cycles: A theoretical, historical, and statistical analysis of the capitalist process* (Vol. 2). McGraw-Hill Book Company.

Syverson, **C.** (2011). What determines productivity? *Journal of Economic Literature*.

Tambe, **P.** (2025). Reskilling the workforce for ai: Domain expertise and algorithmic literacy. *SSRN Electronic Journal*.

Violante, **G. L.** (2002). Technological acceleration, skill transferability, and the rise in residual inequality. *The Quarterly Journal of Economics*.

Appendix

A Model appendix

A.1 Distributional dynamics

Denote lower case variables as the marginal distribution counterparts to their corresponding CDFs. The distribution over matches evolves in the interior of the state space[44]

$$
\begin{aligned}
\partial_t g_t^e(z,h \mid \tau) = &\underbrace{\mathcal{L}^*\left[g_t^e\right](z,h \mid \tau)}_{\text{Learning}} + \underbrace{\mathcal{Q}^*\left[g_t^e\right](z,h \mid \tau)}_{\text{Productivity shocks}} + \underbrace{\mathcal{M}\left[g_t^e\right](z,h \mid \tau; \boldsymbol{g}_t, t)}_{\text{Hiring \& poaching flows}} \\
&+ \underbrace{m_t^a(h)\delta(\tau-1)\delta(z-\underline{z}_t(h))}_{\text{Adoption}} - \underbrace{\kappa g_t^e(z,h \mid \tau)}_{\text{Separations}},
\end{aligned}
\tag{A.1}
$$

where $\boldsymbol{g}_t = \{g_t^e, g_t^v, g_t^u\}$ and $m_t^a(h)$ solves the flux at the boundary $(\underline{z}_t(h), h)$, i.e. the rate at which matched entrepreneurs hit the adoption threshold $\underline{z}_t(h)$[45]

$$
\begin{aligned}
m_t^a(h) = &\left(\alpha\left[h-\underline{z}_t(h)\right]^+ + \vartheta\left(\mu - \underline{z}_t(h)\right)\underline{z}_t(h)\right) g_t^e(\underline{z}_t(h), h \mid 0) - \frac{1}{2}\partial_z\left(\sigma^2 \underline{z}_t(h)^2 g_t^e(\underline{z}_t(h), h \mid 0)\right) \\
&- \underline{z}_t'(h)\alpha\left[\underline{z}_t(h) - h\right]^+ g_t^e(\underline{z}_t(h), h \mid 0).
\end{aligned}
\tag{A.2}
$$

The adjoint operators are given by

$$
\mathcal{L}^*\left[g_t^e\right](z,h \mid \tau) = -\partial_z\left(\alpha_1\left[h-z\right]^+ g_t^e(z,h \mid \tau)\right) - \partial_h\left(\left(\alpha_0 + \alpha_1\left[z-h\right]^+\right) g_t^e(z,h \mid \tau)\right) \tag{A.3}
$$

$$
\mathcal{Q}^*\left[g_t^e\right](z,h \mid \tau) = -\partial_z\left(\left(\vartheta(\mu - z)z\right) g_t^e(z,h \mid \tau)\right) + \frac{1}{2}\partial_{zz}\left(\sigma^2 z^2 g_t^e(z,h \mid \tau)\right) \tag{A.4}
$$

[44] That is for all $(z,h) \in \mathbb{R}^+ \times \mathbb{R}^+$ such that $z \geq \underline{z}_t(h)$. Note that this assumes that such a threshold exists. This is confirmed numerically in the calibrated version of the model.

[45] The probability flux through the boundary $\underline{z}_t(h)$ is given by

$$
R(h,t) = J(\underline{z}_t(h), h \mid 0) \cdot \mathbf{n}(\underline{z}_t(h), h) ds,
$$

where J denotes the probability flux vector at $(\underline{z}_t(h), h)$ from the KFE, $\mathbf{n} = \frac{(-\underline{z}_t'(h),1)}{\sqrt{1+\underline{z}_t'(h)}}$ the unit normal vector to the boundary at $(\underline{z}_t(h), h)$, and $ds = \sqrt{1+\underline{z}_t'(h)^2}dh$ the differential line element along the boundary.

$$\mathcal{M}^*\left[g_t^e\right](z,h\mid\tau;\mathbf{g}_t,t) = \underbrace{-\xi p(\theta_t)\sum_{\tau'\in\mathcal{T}_t}\int \mathbb{1}^{EE}(z,h,z'\mid\tau,\tau';t)\left(\frac{v_t(z'\mid\tau)}{\mathbb{V}_t}g_t^v(z'\mid\tau')\right)dz' g_t^e(z,h\mid\tau)}_{\text{Outflows due to E-E moves}}$$
$$+\underbrace{\xi p(\theta_t)\sum_{\tau_f,\tau_w\in\mathcal{T}_t}\iiint \mathcal{K}^{EE}((h',\tau_w,z',\tau_f),(h,\tau,z,\tau)\mid z'';t)g_t^e(z'',h'\mid\tau_w)dz''dh'\frac{v_t(z'\mid\tau_f)}{\mathbb{V}_t}g_t^v(z'\mid\tau_f)dz'}_{\text{Inflows due to E-E moves}}$$
$$+\underbrace{p(\theta_t)\sum_{\tau_f,\tau_w\in\mathcal{T}_t}\iiint \mathcal{K}^{UE}((h',\tau_w,z',\tau_f),(h,\tau,z,\tau);t)g_t^u(h'\mid\tau_w)dh'\frac{v_t(z'\mid\tau_f)}{\mathbb{V}_t}g_t^v(z'\mid\tau_f)dz'}_{\text{Inflows due to U-E moves}}. \tag{A.5}$$

The transition kernels $\mathcal{K}^{EE}, \mathcal{K}^{UE}$ capture, conditional on meeting a worker, (1) whether a new match is formed (2) which technology vintage τ worker and firm agree to use and (3) the stochastic changes to the productivity vector due to the adoption decision. The kernel therefore captures the joint adoption and hiring decisions of worker-firm pairs as well as the stochastic changes to their productivity vector. The kernels are given by, letting δ denote the Dirac delta function,[46]

$$\mathcal{K}^{EE} = \begin{cases} \delta(h'-h)\delta(z'-z)\mathbb{1}\{\tau=0\} & \text{if } \mathbb{1}^{EE}(z'',h',z'\mid\tau_w,\tau_f;t)=1 \text{ and } \mathbb{1}^{A}(z',h'\mid\tau_f,\tau_w)=0 \\ \mathcal{P}(\tau_w,\tau_f)\delta(h'-\overline{\delta}^{1-\tau_w}h)\delta(z'-\overline{\delta}^{1-\tau_f}z)\mathbb{1}\{\tau=1\} & \text{if } \mathbb{1}^{EE}(z'',h',z'\mid\tau_w,\tau_f;t)=1 \text{ and } \mathbb{1}^{A}(z',h'\mid\tau_f,\tau_w;t)=1 \\ \left(1-\mathcal{P}(\tau_w,\tau_f)\right)\delta(h'-\underline{\delta}^{1-\tau_w}h)\delta(z'-\underline{\delta}^{1-\tau_f}z)\mathbb{1}\{\tau=1\} & \text{if } \mathbb{1}^{EE}(z'',h',z'\mid\tau_w,\tau_f;t)=1 \text{ and } \mathbb{1}^{A}(z',h'\mid\tau_f,\tau_w;t)=1 \\ 0 & \text{if } \mathbb{1}^{EE}(z'',h',z'\mid\tau_w,\tau_f;t)=0, \end{cases} \tag{A.6}$$

and

$$\mathcal{K}^{UE} = \begin{cases} \delta(h'-h)\delta(z'-z)\mathbb{1}\{\tau=0\} & \text{if } \mathbb{1}^{UE}(h',z'\mid\tau_w,\tau_f;t)=1 \text{ and } \mathbb{1}^{A}(z',h'\mid\tau_f,\tau_w;t)=0 \\ \mathcal{P}(\tau_w,\tau_f)\delta(h'-\overline{\delta}^{1-\tau_w}h)\delta(z'-\overline{\delta}^{1-\tau_f}z)\mathbb{1}\{\tau=1\} & \text{if } \mathbb{1}^{UE}(h',z'\mid\tau_w,\tau_f;t)=1 \text{ and } \mathbb{1}^{A}(z',h'\mid\tau_f,\tau_w;t)=1 \\ \left(1-\mathcal{P}(\tau_w,\tau_f)\right)\delta(h'-\underline{\delta}^{1-\tau_w}h)\delta(z'-\underline{\delta}^{1-\tau_f}z)\mathbb{1}\{\tau=1\} & \text{if } \mathbb{1}^{UE}(h',z'\mid\tau_w,\tau_f;t)=1 \text{ and } \mathbb{1}^{A}(z',h'\mid\tau_f,\tau_w;t)=1 \\ 0 & \text{if } \mathbb{1}^{UE}(h',z'\mid\tau_w,\tau_f;t)=0. \end{cases} \tag{A.7}$$

The distribution of solo entrepreneurs evolves as

$$\partial_t g_t^v(z\mid\tau) = \underbrace{\mathcal{Q}^*\left[g_t^v\right](z\mid\tau)}_{\text{Productivity shocks}} + \underbrace{\mathcal{H}^*\left[g_t^v\right](z\mid\tau;\boldsymbol{g}_t,t)}_{\text{Net hiring flows}} + \underbrace{\kappa\int g_t^e(z,h'\tau\mid\tau)dh'}_{\text{Inflows due to separations}}. \tag{A.8}$$

[46]The Dirac delta function is a generalized function such that $\int_{-\infty}^{+\infty}\delta(x)dx = 1$.

The adjoint operator for net hires $\mathcal{H}^*$ is

$$\begin{aligned}\mathcal{H}^*\left[g_t^v\right](z \mid \tau; \boldsymbol{g}_t, t) = &- \underbrace{\xi p(\theta_t) \sum_{\tau \in \mathcal{T}_t} \iint \mathbb{1}^{EE}(z', h', z \mid \tau', \tau; t) g_t^e(z', h' \mid \tau') dh' dz' \left(\frac{v_t(z \mid \tau)}{\mathbb{V}_t} g_t^v(z \mid \tau)\right)}_{\text{Outflow due to successful E-E hires}} \\ &- \underbrace{p(\theta_t) \int \mathbb{1}^{UE}(h', z \mid \tau', \tau; t) g_t^u(h' \mid \tau') dh' \left(\frac{v_t(z \mid \tau)}{\mathbb{V}_t} g_t^v(z \mid \tau)\right)}_{\text{Outflows due to successful U-E hires}} \\ &+ \underbrace{\xi p(\theta_t) \sum_{\tau' \in \mathcal{T}_t} \int \mathbb{1}^{EE}(z, h, z' \mid \tau, \tau'; t) \left(\frac{v_t(z' \mid \tau')}{\mathbb{V}_t} g_t^v(z' \mid \tau')\right) dz' g_t^e(z, h \mid \tau) dh}_{\text{Inflows due to worker getting poached}}.\end{aligned} \tag{A.9}$$

The evolution of the distribution of the unemployed is

$$\begin{aligned}\partial_t g_t^u(h \mid \tau) = &\underbrace{-\alpha_u \partial_h g_t^u(h \mid \tau)}_{\text{Human capital depreciation}} + \underbrace{\kappa \int g_t^e(z', h \mid \tau) dz'}_{\text{Inflows due to separations}} \\ &\underbrace{-p(\theta_t) \sum_{\tau' \in \mathcal{T}_t} \int \mathbb{1}^{UE}(h, z' \mid \tau, \tau'; t) \frac{v_t(z' \mid \tau')}{\mathbb{V}_t} g_t^v(z' \mid \tau') dz' g_t^u(h \mid \tau)}_{\text{Outflows due to U-E move}}.\end{aligned} \tag{A.10}$$

A.2 Value function operators

I define the following operators which will become useful in the subsequent sections.

Learning operator. Let

$$\mathcal{L}[S_t](z, h \mid \tau) = \partial_z S_t(z, h \mid \tau) \alpha \left[h - z\right]^+ + \partial_h S_t(z, h \mid \tau) \alpha \left[z - h\right]^+$$

capture changes in the value function due to the mutual worker-firm learning.

Productivity operator. Let shocks to productivity be captured by

$$\mathcal{Q}\left[S_t\right](z, h \mid \tau) = \partial_z S_t(z, h \mid \tau) \vartheta(\mu - \log z) z + \frac{\sigma^2 z^2}{2} \partial_{zz} S_t(z, h \mid \tau).$$

With some abuse of notation, let

$$\mathcal{Q}\left[V_t\right](z, h \mid \tau) = \mathcal{Q}\left[V_t\right](z \mid \tau) = \partial_z V_t(z \mid \tau) \vartheta(\mu - \log z) z + \frac{\sigma^2 z^2}{2} \partial_{zz} V_t(z \mid \tau).$$

Vacancy posting gain. Let the vacancy posting gain to solo entrepreneurs be given by

$$\mathcal{G}_t(z \mid \tau) = \max\left\{ vq(\theta_t)\frac{1}{\mathbb{X}_t}\left[\xi \sum_{\tau'\in\mathcal{T}_t}\int \Delta_t^{EE}(z',h',z|\tau',\tau)^+ dG_t^e(z',h'|\tau') + \sum_{\tau'\in\mathcal{T}_t}\int \Delta_t^{UE}(h',z|\tau',\tau)^+ dG_t^u(h'|\tau')\right] - c(v)\right\}.$$

Note that this operator is non-linear.

A.3 Poaching policies

By Equation (3.5) a worker is poached whenever $\mathbb{1}^{EE} = 1$. For illustrative purposes, let us consider the case without adoption. Then, poaching is successful whenever

$$S_t(z,h \mid 0) - S_t(z',h \mid 0) > V_t(z \mid 0) - V_t(z' \mid 0).$$

To understand the forces behind this equation let us consider each side of the equation in turn. The left-hand side satisfies

$$\begin{aligned}
S_t(z,h \mid 0) - S_t(z',h \mid 0) = & \underbrace{\frac{\pi(z,h \mid 0) - \pi(z',h \mid 0)}{\rho+\kappa}}_{\text{static production}} \\
& + \underbrace{\frac{\mathcal{L}\left[S_t\right](z,h \mid 0) - \mathcal{L}\left[S_t\right](z',h \mid 0)}{\rho+\kappa}}_{\text{differential learning}} \\
& + \underbrace{\frac{\mathcal{Q}\left[S_t\right](z,h \mid 0) - \mathcal{Q}\left[S_t\right](z',h \mid 0)}{\rho+\kappa}}_{\text{differential exogenous shocks}} \\
& + \underbrace{\frac{\kappa}{\kappa+\rho}\left(V_t(z \mid 0) - V_t(z' \mid 0)\right)}_{\text{differential outside options}} \\
& + \underbrace{\frac{\partial_t S_t(z,h \mid 0) - \partial_t S_t(z',h \mid 0)}{\rho+\kappa}}_{\text{differential capital gains}}.
\end{aligned}$$

The right-hand side satisfies

$$V_t(z \mid 0) - V_t(z' \mid 0) = \underbrace{\frac{\pi(z,0 \mid 0) - \pi(z',0 \mid 0)}{\rho + d}}_{\text{differential production}} + \underbrace{\frac{\mathcal{Q}[V_t](z \mid 0) - \mathcal{Q}[V_t](z' \mid 0)}{\rho + d}}_{\text{differential exogenous shocks}} + \underbrace{\frac{\mathcal{G}_t(z \mid 0) - \mathcal{G}_t(z' \mid 0)}{\rho + d}}_{\text{differential hiring gains}} + \underbrace{\frac{\partial_t V_t(z \mid 0) - \partial_t V_t(z' \mid 0)}{\rho + d}}_{\text{differential capital gains}}.$$

Noting that $\frac{\kappa}{\kappa+\rho} - 1 = -\frac{\rho}{\kappa+\rho}$, this delivers the following relation determining whether a worker is successfully poached

$$\begin{aligned}\left(\pi(z,h \mid 0) - \frac{\rho}{\rho+d}\pi(z,0 \mid 0)\right) - \left(\pi(z',h \mid 0) - \frac{\rho}{\rho+d}\pi(z',0 \mid 0)\right)& \\ + \mathcal{L}\left[S_t\right](z,h \mid 0) - \mathcal{L}\left[S_t\right](z',h \mid 0)& \\ + \mathcal{Q}\left[S_t - \frac{\rho}{\rho+d}V_t\right](z,h \mid 0) - \mathcal{Q}\left[S_t - \frac{\rho}{\rho+d}V_t\right](z',h \mid 0)& \\ + \frac{\rho}{\rho+d}\left(\mathcal{G}_t(z' \mid 0) - \mathcal{G}_t(z \mid 0)\right)& \\ + \partial_t\left(S_t(z,h \mid 0) - \frac{\rho}{\rho+d}V_t(z \mid 0)\right) - \partial_t\left(S_t(z',h \mid 0) - \frac{\rho}{\rho+d}V_t(z' \mid 0)\right) &> 0.\end{aligned}$$

Interpretation. To understand these expressions it is instructive to consider limiting cases. First, suppose that $\zeta \to 1$ and $d \to 0$ such that $\pi(z,h \mid 0) = z + h$. In that case, there are no production complementarities and the expression in the first line is zero. That is, worker mobility is primarily pinned down by differential learning gains (second line) and differential abilities to hire a new worker (fourth line).[47] Another force operates through differential exposure to productivity shocks. As ζ falls production complementarities become the more dominant force while for $d > 0$ more productive entrepreneurs have greater incentives avoiding the riskier outside option of being solo. That is, by becoming solo they would face the risk of exiting and receiving a value of zero, which is more costly to more productive entrepreneurs.

In another extreme, suppose that $z' = h$. In that case $\mathcal{L}\left[S_t\right](z',h \mid 0) = 0$ and there are no learning gains

[47] This in turn is pinned down by the offer distribution of marginal values to the entrepreneur, i.e. the distribution of positive Δ^{EE} and Δ^{UE}.

available. We require $z \neq h$ for there to be positive learning gains.

Introducing adoption. The decision whether to poach or not now depends on — letting $\tau_f = 0$ but $\tau_w = 1$.

$$\max\{S_t^A(z,h \mid 0,1) - S_t(z,h \mid 0), 0\} + S_t(z,h \mid 0) - S_t(z,h \mid 1) > V_t(z \mid 1) - V_t(z \mid 1).$$

In total

$$\begin{aligned}
(\rho + \kappa) &\max\{S_t^A(z,h \mid 0,1) - S_t(z,h \mid 0), 0\} \\
&+ \left(\pi(z,h \mid 0) - \frac{\rho}{\rho+d}\pi(z,0 \mid 0)\right) - \left(\pi(z',h \mid 1) - \frac{\rho}{\rho+d}\pi(z',0 \mid 1)\right) \\
&+ \mathcal{L}\left[S_t\right](z,h \mid 0) - \mathcal{L}\left[S_t\right](z',h \mid 1) \\
&+ \mathcal{Q}\left[S_t - \frac{\rho}{\rho+d}V_t\right](z,h \mid 0) - \mathcal{Q}\left[S_t - \frac{\rho}{\rho+d}V_t\right](z',h \mid 1) \\
&+ \frac{\rho}{\rho+d}\left(\mathcal{G}_t(z' \mid 1) - \mathcal{G}_t(z \mid 0)\right) \\
&+ \partial_t\left(S_t(z,h \mid 0) - \frac{\rho}{\rho+d}V_t(z \mid 0)\right) - \partial_t\left(S_t(z',h \mid 1) - \frac{\rho}{\rho+d}V_t(z' \mid 1)\right) > 0.
\end{aligned}$$

There are two opposing forces: (1) there are additional gains from poaching due to the option value to adopt, captured by $(\rho + \kappa)\max\{S_t^A(z,h \mid 0,1) - S_t(z,h \mid 0), 0\}$, but (2) it becomes relatively harder to poach since the marginal value of the worker to a matched entrepreneur now is scaled. This is seen in a straightforward manner for the limit case in which $\zeta = 1$ and $\pi(z,h \mid 1) - \pi(z,0 \mid 1) = A(z+h) - k_f - (Az - k_f) = Ah$, which scales in A. Further, in the limit as $\xi \to 1$ we get that

$$\pi(z,h \mid 0) - \frac{\rho}{\rho+d}\pi(z,0 \mid 0) - \left(\pi(z',h \mid 1) - \frac{\rho}{\rho+d}\pi(z',0 \mid 1)\right) = -(A-1)h - \frac{d}{\rho+d}\left(Az' - k_f - z\right).$$

A.4 Hiring policies

By Equation (3.5) a worker is poached whenever $\mathbb{1}^{UE} = 1$. For illustrative purposes let us consider the case without adoption. Then, hiring is successful if and only if

$$S_t(z,h \mid 0) - V_t(z \mid 0) - U_t(h \mid 0) > 0.$$

That is, if the match generates a positive surplus.

The left-hand side is given by

$$
\begin{aligned}
(\rho+\kappa)\left(S_t(z,h\mid 0)-V_t(z\mid 0)-U_t(z\mid 0)\right)= & \underbrace{\pi(z,h\mid 0)-\frac{\rho}{\rho+d}\pi(z,0\mid 0)-bh}_{\text{static surplus}} \\
& +\underbrace{\mathcal{L}\left[S_t\right](z,h\mid 0)}_{\text{Learning}} \\
& -\underbrace{\left(-\frac{\alpha_u}{\rho+\alpha_u}bh\right)}_{\text{Human capital depreciation}} \\
& +\underbrace{\mathcal{Q}\left[S_t\right](z,h\mid 0)-\frac{\rho}{\rho+d}\mathcal{Q}\left[V_t\right](z\mid 0)}_{\text{exogenous shocks}} \\
& -\underbrace{\frac{\rho}{\rho+d}\mathcal{G}_t(z\mid 0)}_{\text{option value of hiring}} \\
& +\underbrace{\partial_t\left(S_t(z,h\mid 0)-\frac{\rho}{\rho+d}V_t(z\mid 0)\right)}_{\text{capital gains}} \\
\geq & 0.
\end{aligned}
$$

In the limit of no production complementarities the first line becomes $(1-b)h$, which is positive as long as $b<1$. The second line denotes the learning gains. The third line denotes the cost of unemployment through human capital depreciation. The fourth line captures the exogenous shocks. The forth line denotes the opportunity cost as the solo entrepreneur foregoes the option value of hiring a different worker.

A.5 Adoption decision

A.5.1 Internal adoption

For simplicity let $\delta = \bar{\delta} = \underline{\delta}$. Then,

$$\begin{aligned} S_t(\delta z, \delta h \mid 1) - S_t(z, h \mid 0) = & \underbrace{\frac{\pi(\delta z, \delta h \mid 1) - \pi(z, h \mid 0)}{\rho + \kappa}}_{\text{Short-run output differences}} \\ & + \underbrace{\frac{\mathcal{L}\left[S_t\right](\delta z, \delta h \mid 1) - \mathcal{L}\left[S_t\right](z, h \mid 0)}{\rho + \kappa}}_{\text{Differential learning gains}} \\ & + \underbrace{\frac{\mathcal{Q}\left[S_t\right](\delta z, \delta h \mid 1) - \mathcal{Q}\left[S_t\right](z, h \mid 0)}{\rho + \kappa}}_{\text{Differential productivity shocks}} \\ & + \underbrace{\frac{\kappa}{\rho + \kappa}\left(V_t(\delta z \mid 1) - V_t(z \mid 0)\right)}_{\text{Differential changes in outside value for entrepreneur}} \\ & + \underbrace{\frac{\kappa}{\rho + \kappa}\left(U_t(\delta h \mid 1) - U_t(h \mid 0)\right)}_{\text{Differential changes in outside value for entrepreneur}} \quad . \end{aligned} \tag{A.11}$$

Note that $\pi(\delta z, \delta h \mid 1) - \pi(z, h \mid 0) = (A\delta - 1)\,\pi(z, h \mid 0) - k_f$. If this is positive, then adoption is immediately profitable. If not, then adoption may still occur through the joint learning and productivity shock process. First, mean reversion implies a form of exogenous learning whenever $\log z$ falls below the long-run mean. Second, the volatility harvesting process implies that match duration induces a positive trend in z and h jointly. This may make adoption attractive even if it delivers short-run losses. Third, if an entrepreneur adopts with a highly-skilled worker, i.e. large h, then this learning gain persists in the new technology and increase the trend growth of the entrepreneur's productivity.

The fourth line captures the trade-off between (1) insurance from wide-spread adoption and (2) free-riding mechanism. When the pool of adopters is large then it is easier to find a previous adopter when solo and can immediately form more productive matches relative to the counterfactual of a small pool of such workers. That is, these forces increase $V_t(\delta z \mid 1)$. On the other hand, a large pool of such workers also increases the opportunity of adoption via hiring which is less costly to the entrepreneur. These forces are traded off as $V_t(\delta z \mid 1) - V_t(z \mid 0)$ and discounted by the likelihood of an exogenous separation.

Further, the only way aggregate conditions matter here is through the outside option V_t. The presence of long-lasting vacancies, i.e. vacancy chains, is how the result in Lise and Robin (2017) that surplus is independent of aggregate distributions is broken.

A.5.2 Hire to adopt

Let us now consider when a new match with initial states (z, τ_f) and (h, τ_w) decides to use the new technology when $\tau_f = 0$ and $\tau_w = 1$. Then,

$$
\begin{aligned}
S_t(\delta z, h \mid 1) - S_t(z, h \mid 0) = & \underbrace{\frac{\pi(\delta z, h \mid 1) - \pi(z, h \mid 0)}{\rho + \kappa}}_{\text{Short-run output differences}} \\
& + \underbrace{\frac{\mathcal{L}\left[S_t\right](\delta z, h \mid 1) - \mathcal{L}\left[S_t\right](z, h \mid 0)}{\rho + \kappa}}_{\text{Differential learning gains}} \\
& + \underbrace{\frac{\mathcal{Q}\left[S_t\right](\delta z, h \mid 1) - \mathcal{Q}\left[S_t\right](z, h \mid 0)}{\rho + \kappa}}_{\text{Differential productivity shocks}} \\
& \underbrace{+ \frac{\kappa}{\rho + \kappa}\left(V_t(\delta z \mid 1) - V_t(z \mid 0)\right)}_{\text{Differential changes in outside value for entrepreneur}} .
\end{aligned}
\tag{A.12}
$$

The major difference relative to above is that since the worker has already accumulated technology-specific expertise only the entrepreneur is subject to the limited transferability. As a result, learning gains open up which further incentivize the adoption of the technology — as captured by the second line.

A.6 Decomposition of new adopters

A.6.1 Newly adopting entrepreneurs

Let the mass of entrepreneurs using the new technology be denoted by μ_t^E and the share by μ_t^E/M. Then,

$$
\begin{aligned}
\dot{\mu}_t^E = & \underbrace{\int m_t^a(h)\delta(z - \underline{z}_t(h))dhdz}_{\text{internal adoption}} \\
& \underbrace{+ \xi p(\theta_t) \sum_{\tau\in\{0,1\}} \int \frac{v_t(z \mid 0)}{\mathbb{V}_t} \int \mathbb{1}^{EE}\left(z', h', z \mid \tau, 0; t\right) \mathbb{1}^{A}\left(h', z \mid \tau, 0; t\right) dG_t^e(z', h' \mid \tau) dG_t^v(z \mid 0)}_{\text{Poach to adopt}} \\
& \underbrace{+ p(\theta_t) \sum_{\tau\in\{0,1\}} \int \frac{v_t(z \mid 0)}{\mathbb{V}_t} \int \mathbb{1}^{UE}\left(h', z \mid \tau, 0; t\right) \mathbb{1}^{A}\left(h', z \mid \tau, 0; t\right) dG_t^u(h' \mid \tau) dG_t^v(z \mid 0)}_{\text{Hire unemp. to adopt}} \\
& \underbrace{- \xi p(\theta_t) \sum_{\tau\in\{0,1\}} \int \frac{v_t(z \mid 0)}{\mathbb{V}_t} \int \mathbb{1}^{EE}\left(z', h', z \mid \tau, 0; t\right) \left(1 - \mathbb{1}^{A}\left(h', z \mid \tau, 0; t\right)\right) dG_t^e(z', h' \mid \tau) dG_t^v(z \mid 1)}_{\text{Downgrade technology}} \\
& \underbrace{- p(\theta_t) \sum_{\tau\in\{0,1\}} \int \frac{v_t(z \mid 0)}{\mathbb{V}_t} \int \mathbb{1}^{UE}\left(h', z \mid \tau, 0; t\right) \left(1 - \mathbb{1}^{A}\left(h', z \mid \tau, 0; t\right)\right) dG_t^u(h' \mid \tau) dG_t^v(z \mid 1)}_{\text{Downgrade technology}}
\end{aligned}
\tag{A.13}
$$

A.6.2 Newly adopting workers

Let the mass of workers using the new technology be denoted by μ_t^W which is equivalent to its share since the population of workers is normalized to one. Then,

$$\begin{aligned}
\dot{\mu}_t^W = & \underbrace{\int m_t^a(h)\delta(z - \underline{z}_t(h))dhdz}_{\text{internal adoption}} \\
& \underbrace{+ \xi p(\theta_t) \sum_{\tau \in \{0,1\}} \int \frac{\upsilon_t(z \mid \tau)}{\mathbb{V}_t} \int \mathbb{1}^{EE}\left(z', h', z \mid 0, \tau; t\right) \mathbb{1}^{A}\left(h', z \mid 0, \tau; t\right) dG_t^e(z', h' \mid 0) dG_t^\upsilon(z \mid \tau)}_{\text{Poached \& adopt}} \\
& \underbrace{+ p(\theta_t) \sum_{\tau \in \{0,1\}} \int \frac{\upsilon_t(z \mid \tau)}{\mathbb{V}_t} \int \mathbb{1}^{UE}\left(h', z \mid 0, \tau; t\right) \mathbb{1}^{A}\left(h', z \mid 0, \tau; t\right) dG_t^u(h' \mid \tau) dG_t^\upsilon(z \mid 0)}_{\text{Hired \& adopt}} \\
& \underbrace{- \xi p(\theta_t) \sum_{\tau \in \{0,1\}} \int \frac{\upsilon_t(z \mid \tau)}{\mathbb{V}_t} \int \mathbb{1}^{EE}\left(z', h', z \mid 1, \tau; t\right)\left(1 - \mathbb{1}^{A}\left(h', z \mid 1, \tau; t\right)\right) dG_t^e(z', h' \mid 1) dG_t^\upsilon(z \mid \tau)}_{\text{Downgrade technology}} \\
& \underbrace{- \xi p(\theta_t) \sum_{\tau \in \{0,1\}} \int \frac{\upsilon_t(z \mid \tau)}{\mathbb{V}_t} \int \mathbb{1}^{UE}\left(h', z \mid 1, \tau; t\right)\left(1 - \mathbb{1}^{A}\left(h', z \mid 1, \tau; t\right)\right) dG_t^u(h' \mid 1) dG_t^\upsilon(z \mid \tau)}_{\text{Downgrade technology}}
\end{aligned} \tag{A.14}$$

A.7 Further details on learning dynamics

In this section, I show that for matches that persist forever, h and z grow unboundedly. Growth comes from systematically capturing upside volatility. A heuristic argument can be made as follows. Solving the SDE

$$\mathbb{E}_t\left[\log h(t)\right] = \log h(0) + \alpha \int_0^t \left[z(s)/h(s) - 1\right]^+ ds. \tag{A.15}$$

Since $dz(t)$ is an Ornstein-Uhlenbeck process, for any scalar $L \in \mathbb{R}$ $z(s)$ almost surely crosses L infinitely times. Hence, for any $L \geq h(s)$ there exists some T such that $z(T) > h(s)$. The worker learns from the firm and h grows. Since z moves continuously these crossings contribute positive areas and these non-zero contributions make $h(t) \to \infty$. Then, since $h(t)$ grows unboundedly the following will as well since

$$\mathbb{E}_t\left[\log z(t)\right] = e^{-\vartheta t} \log z(0) + \alpha \int_0^t e^{-\vartheta(t-s)} \left[h(s)/z(s) - 1\right]^+ ds. \tag{A.16}$$

A.8 Special case: Long-run

Let the vacancy posting cost and the unemployment benefit b scale in aggregate output Y_t. Let $d = 0$. Then,

$$\lim_{t\to\infty} S_t(z, h \mid 1) = AS_{ss}(z, h \mid 0) - k_f/\rho \tag{A.17}$$

and

$$\lim_{t\to\infty} V_t(z, h \mid 1) = AV_{ss}(z, h \mid 0) - k_f/\rho \tag{A.18}$$

and

$$\lim_{t\to\infty} U_t(h \mid 1) = AU_{ss}(h \mid 0) \tag{A.19}$$

All decisions are unchanged, and therefore the same distribution exists. So can chain these together giving long-run growth through these waves.

B Calibration Appendix

We map our empirical estimate to the average probability of an employed worker moving to a newly adoption firm in period t. This is informative of degree to which poaching is related to adoption. Formally, it is given by

$$\int \left[\xi p(\theta_t) \left(\frac{v_t(z \mid 0)}{\mathbb{V}_t} \right) \int \mathbb{1}^{EE}(z', h, z \mid 1, 0) \mathbb{1}^{A}(z, h \mid 0, 1) \frac{dG^e_t(z', h \mid 1)}{\int dG^e_t(z', h \mid 1)} \right] \frac{dG^v_t(z \mid 0)}{\int dG^v_t(z \mid 0)}.$$

We show the time path of this transition in Figure B.1, which is similar to our empirical estimates for senior workers in the early stages of the transition.

The full calibration table is shown below

Table B.1: Calibrated parameters

Parameter	Description	Value	Comment
		General parameters	
ρ	discount rate	0.02	8% annual
M	mass of entrepreneurs	1.2	-
d	exit rate (solo entrepreneurs)	$\varkappa$	-
		Labor market parameters	
χ	matching efficiency	0.725	unemployment rate of 5%
β	matching elasticity	0.5	standard
ξ	relative search intensity (J2J)	0.215	E-E rate
$\varkappa$	exogenous separation rate	0.038	E-U rate
b	benefits	0	-
		Entrepreneur's parameters	
ζ	production complementarities	0.8	Herkenhoff et al. (2024)
σ^2	volatility	0.005	90-10 ratio
ϑ	persistence	0.02	-
μ	long-run mean	$e^{-\sigma^2/(2\vartheta)}$	normalization
$\bar{c}$	vacancy cost scaling	1	normalization
γ_v	convexity of vacancy cost	3.450	Bilal et al. (2022)
		Learning parameters	
α	learning rate	0.01	Herkenhoff et al. (2024)
α_u	human capital depreciation	0.032	Herkenhoff et al. (2024)
		Transition parameters	
A	lower mc	1.5	exercise
k_f	higher fixed cost	0.225	-
$\overline{\delta}$	high transferability	0.915	-
$\underline{\delta}$	low transferability	0.715	-
$\mathcal{P}(0,0)$		0	normalization
$\mathcal{P}(1,0)$		0.4	empirical target
$\mathcal{P}(0,1)$		0.4	symmetry

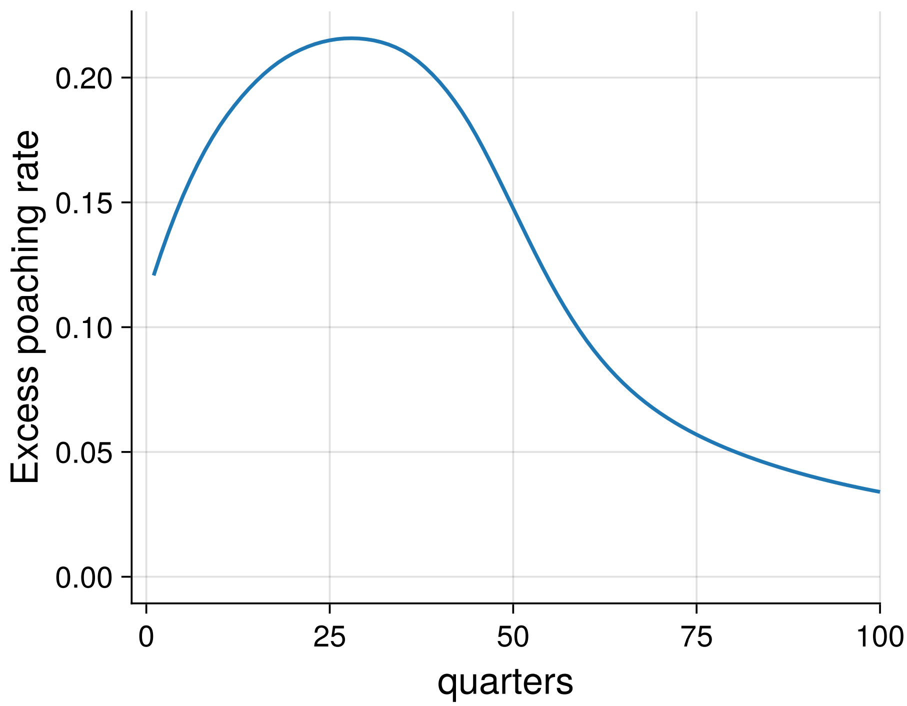


Figure B.1: Probability of moving to non-adopter.

C Empirical Appendix

C.1 Distribution of adoption years

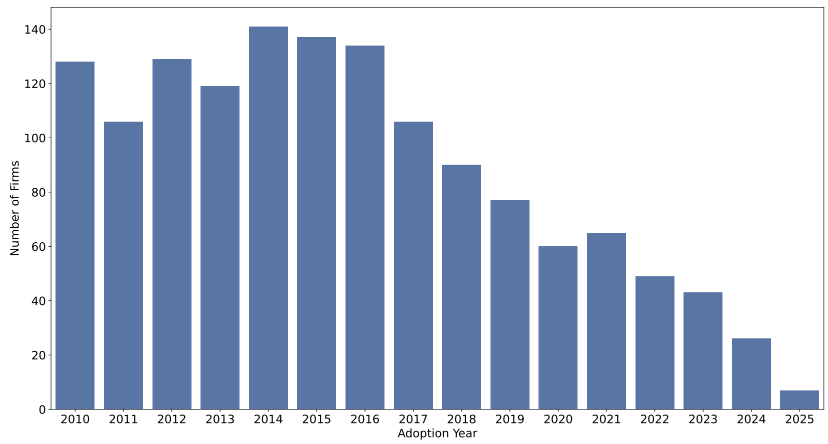


Figure C.2: Distribution of adoption years of Compustat sample (2010-2024)

C.2 Share of external AI hires

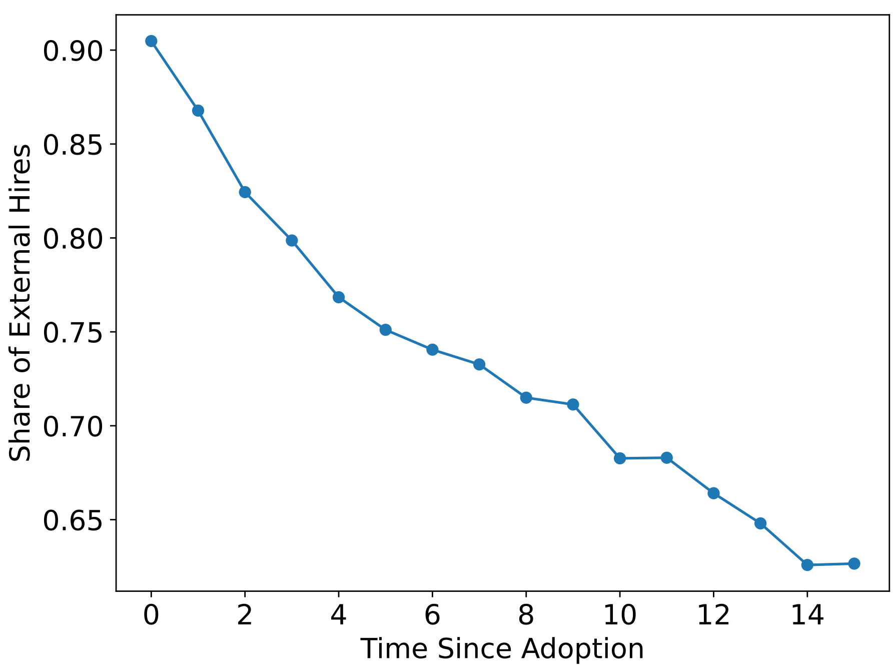


Figure C.3: Share of external hires of AI jobs

C.3 Subsample restrictions.

The following subsample restrictions are applied to the raw data. The main analysis covers years 2010 to 2024, a time period with good coverage and substantial AI adoption. The distribution of AI adoption during this sample is shown in Figure C.2. I restrict the focus to professional workers (SOCs 11-19) in metropolitan MSAs working in professional services, tech, and finance industries and high-tech manufacturing.[48] The final analysis uses worker flows where the destination firm corresponds to a publicly traded firm from Compustat while the origin firm is public or private. This allows us to retain as much information as possible regarding the movement of workers while supplementing the data with additional information on e.g. firm sales.

C.4 Identification argument

How should we think about the correlations documented in Section 2? We will now discuss the identification argument through which we can view the presented correlation. A classic concern in empirical studies of technology adoption is that selection into adoption is non-random and endogenous to the expected return of using the technology. To address this issue, I follow an approach similar to Humlum (2019). Specifically, the regression in (2) includes as controls pre-adoption sales per worker. The crucial assumption is that revenue productivity (growth) and industry-firm fixed effects are good proxies for a firm's expected return of adoption. Then, conditional on those controls, access to highly skilled workers with prior AI experience is a binding constraint for adoption and firms successfully implement AI only if they match with the right type of worker – a process that is at least in part random and an integral feature of canonical labor search models. I use this interpretation to map our empirical estimates to the quantitative model.

C.5 Robustness

The estimates from Section 2 are robust to a series of variations.

C.5.1 Excluding initial AI adopters

A concern is that we use labor market statistics to both identify adoption events and measure worker flows. To mitigate the extent to which this relationship could be mechanical in nature, I next exclude the workers which are used to identify adoption events from the analysis. That is, all AI workers that appear in a firm during the first year that AI is used in the firm are excluded from the analysis for that year. Estimates are largely unchanged.

[48] The included industry codes (NAICS) are 51, 52, 54, 55, 3341, 3342, 3344, 3345, 3364.

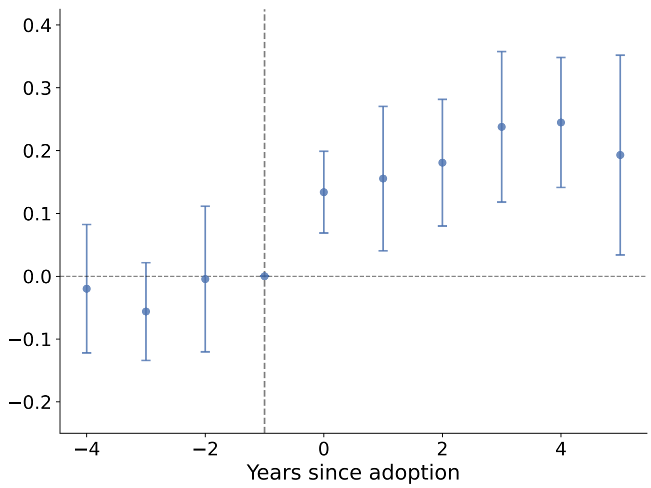


Figure C.4: Excess probabilities. Excluding first AI workers in firm.

Notes: This graph plots the coefficient estimates β_h from (2) and corresponds to how excess probabilities $\gamma_{cj\ell t+h}$ change with distance to the adoption year t. This version excludes first AI adopters in a firm. Estimated on subsample of E-E movers at annual frequency. The regression controls for industry-year and location fixed effects as well as pre-adoption variables related to selection into adoption (e.g. sales per worker). The sample period is 2010-2024. Destination firms in the sample are Compustat firms. Origin firms in the sample is the universe of firms in Revelio Labs. Standard errors are clustered at the firm-year level.

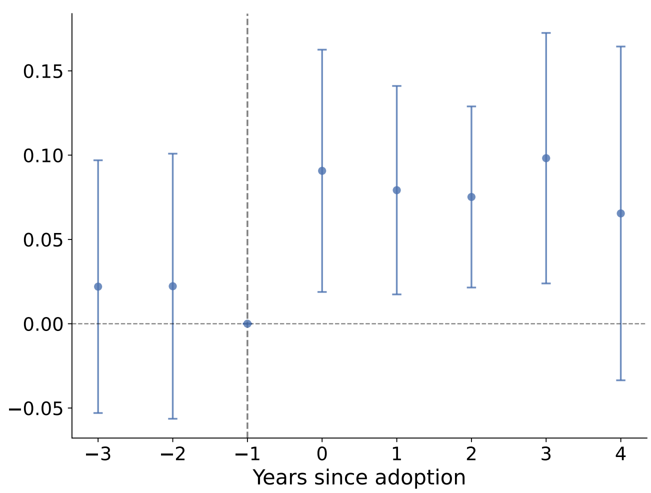
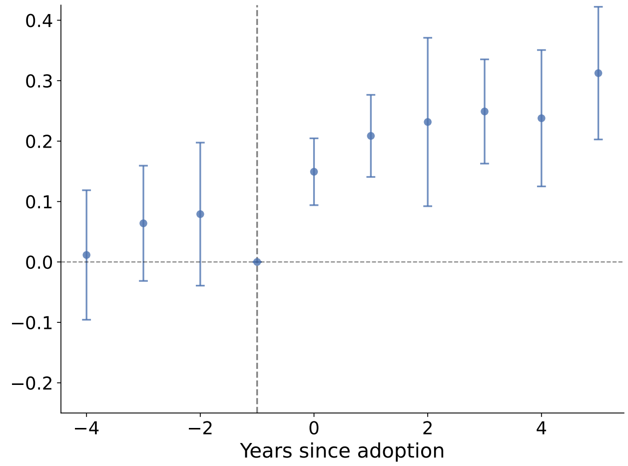
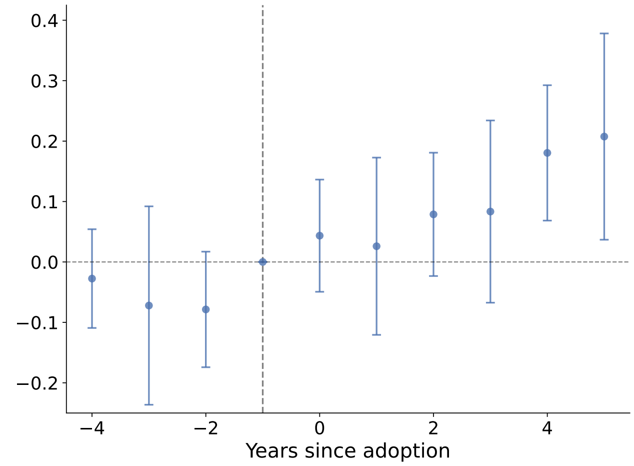


Figure C.5: Excess probabilities. Including firm-level fixed effects. Excluding first AI workers in firm.

Notes: This graph plots the coefficient estimates β_h from (2) and corresponds to how excess probabilities $\gamma_{cj\ell t+h}$ change with distance to the adoption year t. This version excludes first AI adopters in a firm. Estimated on subsample of E-E movers at annual frequency. The regression controls for firm-level, industry-year and location fixed effects as well as pre-adoption variables related to selection into adoption (e.g. sales per worker). The sample period is 2010-2024. Destination firms in the sample are Compustat firms. Origin firms in the sample is the universe of firms in Revelio Labs. Standard errors are clustered at the firm-year level.



(a) Senior



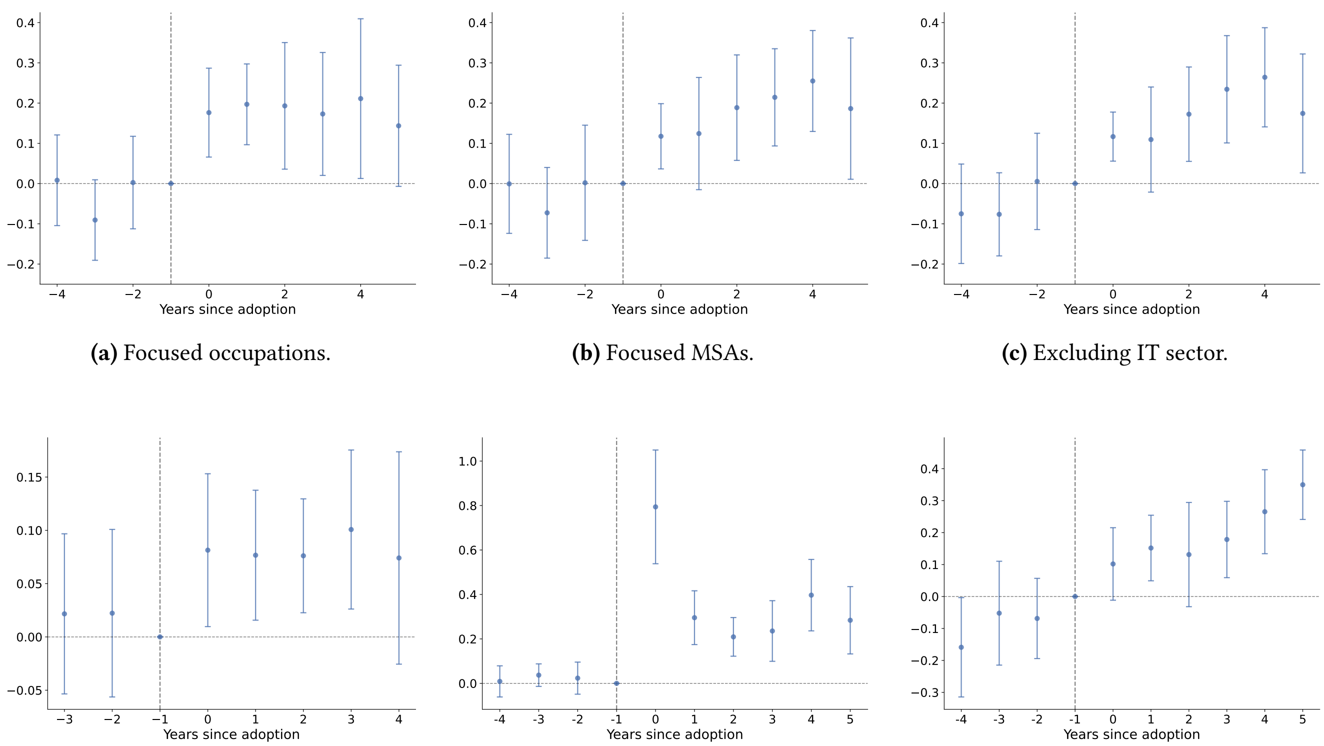
(b) Junior

Figure C.6: Excess probabilities by seniority status. Excluding first AI workers in firm.

Notes: This graph plots the coefficient estimates β_h^c from (2) and corresponds to how excess probabilities $\gamma_{cj\ell t+h}$ change with distance to the adoption year t, separately for two subgroups c: senior workers (e.g. managers and directors) and junior workers (e.g. analysts). Excluding first AI adopters in firm. The left panel shows coefficients for the senior subsample. The right panel shows coefficients for the junior subsample. Estimated on subsample of E-E movers at annual frequency. The regression controls for industry-year and location fixed effects as well as pre-adoption variables related to selection into adoption (e.g. sales per worker). The sample period is 2010-2024. Destination firms in the sample are Compustat firms. Origin firms in the sample is the universe of firms in Revelio Labs. Standard errors are clustered at the firm-year level.

C.5.2 Other specifications

We explore two other specifications of the baseline regression in Section 2.3. In Figure C.7d we show the estimated coefficients of (2) where we additionally include destination firm fixed effects α_j. In Figure C.7e we show the estimated coefficients of a modified version of (1) where c denotes the set of workers from all firms that have previously utilized AI. The indicator D_{ikt} now takes the value one if the worker has previously held an AI position. The estimates indicate that it is really the AI adopters that see the increase in the hiring.



(a) Focused occupations.

(b) Focused MSAs.

(c) Excluding IT sector.

(d) Including firm fixed effects.

(e) AI workers from previous adopters.

(f) Hiring from other industries.

Figure C.7: Alternative specification robustness.

D Additional model results

D.1 Free-riding motives over transition

A defining feature of Figure 10 is that the contribution of internal adoption to newly adopting entrepreneurs is declining over the transition. In fact, by quarter 30 contributions are essentially zero. What drives these dynamics?

Motives for internal adoption are detailed in A.5.1. As discussed, time-varying incentives from adopting the new technology internally stem from changes in $V_t(\delta z \mid 1) - V_t(z \mid 0)$. Figure D.8 plots the adoption thresholds $\underline{z}_t(h)$ up to quarter 40. This illustrates that adoption thresholds are continuously shifting outwards, indicating that the difference $V_t(\delta z \mid 1) - V_t(z \mid 0)$ is falling over time.

The intuition is as follows. As the pool of workers experienced in the new technology grows so does the likelihood that a solo entrepreneur meets such a worker and is able to adopt the new technology via hires — either through poaching or out of unemployment. This generates free-rider motives, i.e. foregoing the more costly internal adoption in favor of adopting via hiring. As a result, the adoption threshold shifts outwards.

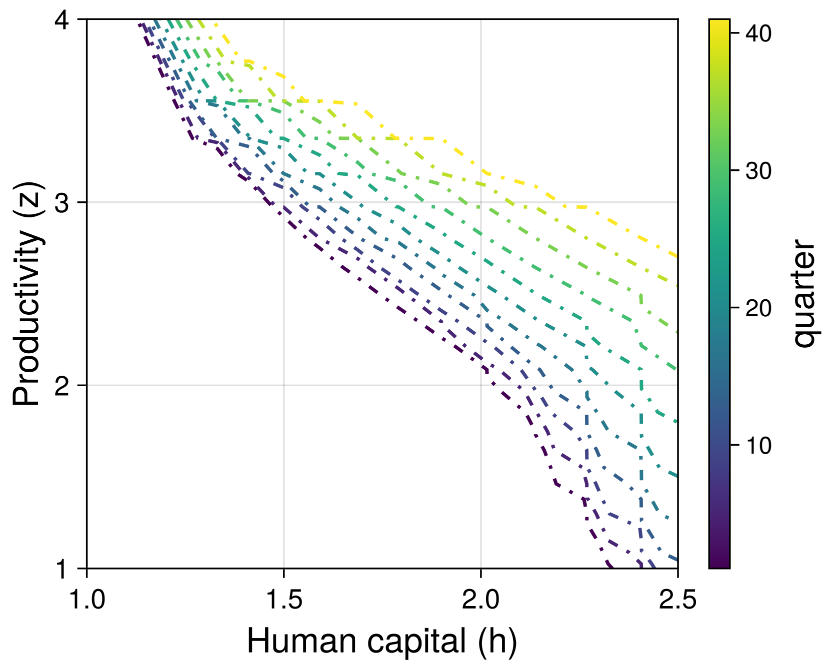


Figure D.8: Adoption thresholds over the transition.

Notes: This plot shows the adoption thresholds $\underline{z}_t(h)$ at different times during the transition and up to quarter 40. Darker, purple lines illustrate early in the transition. Brighter, yellow lines illustrate later in the transition.

However, this may not necessarily be socially inefficient. There are social benefits from adopting with the help of knowledge spillovers of experienced workers and entrepreneurs. In contrast, internal adoption in the presence of a large pool of experienced workers can be wasteful and inefficient. This also underlies the argument developed in Section 5.2 that adoption can be too fast.

D.2 Increased skill supply via education

Consider the following extension. Workers retire at rate ϱ at which point they are replaced by new workers who enter with the same level of human capital h, but may be biased towards the new technology vintage $\tau = 1$. Specifically, a new worker replacing a retiree with human capital h enters as trained in the new technology with density $f_t(h \mid 1)$, where[49]

$$f_t(h \mid 1) = \min\left\{\omega, \frac{1-u_t}{\iint g_t^e(z,h \mid 1)}\right\} \int g_t^e(z,h \mid 1)dz + g_t^u(h \mid 1).$$

The min operator ensures that $f_t(h \mid 0) \geq 0$, but will only be binding towards the end of the transition or for very large values of ω. Let the density for entering with the old technology be given by

$$f_t(h \mid 0) = \frac{\max\left\{1-u_t - \omega \iint g_t^e(z,h \mid 1)dzdh, 0\right\}}{\iint g_t^e(z,h \mid 0)dzdh} \int g_t^e(z,h \mid 0)dz + g_t^u(h \mid 0).$$

It can be verified that indeed $\int f_t(h \mid 0)dh + \int f_t(h \mid 1)dh = 1$, with $f_t(h \mid \tau) \geq 0$ for all $h, \tau \in \mathbb{R}^+ \times \{0,1\}$ and this is a valid density.

Comments.. The above representation captures in a reduced form any factors inducing a slow-moving increase in the supply of workers experienced in the new technology that are increasing in the current stock of such workers. In the main text, this is captured via exogenous separations: When more workers experienced in the new technology are employed, then the unconditional probability of a skilled worker becoming unemployed increases. Since hiring out of unemployment is easier for entrepreneurs this captures the notion of slow moving reallocation mechanism. Similarly, in the formulation above the distribution of entrants is closely linked with the distribution of active workers. This has close parallels with the literature on knowledge diffusion by imitation, but could also capture in a reduced form way endogenous education decisions by workers which, for expositional simplicity, are not modeled explicitly here. For example, greater observed employment opportunities in the new technology for a specific worker h, i.e. a larger $\int g_t^e(z,h \mid 1)dz$, may induce larger incentives to obtain required skills during the educational years.

[49]Note the similarity to models of imitation by entrants as in Sampson (2015).

Distributional dynamics. Adapting the distributional dynamics form (A.10) to the current setting gives

$$\partial_t g_t^u(h \mid \tau) = \underbrace{-\alpha_u \partial_h g_t^u(h \mid \tau)}_{\text{Human capital depreciation}} \underbrace{+\kappa \int g_t^e(z', h \mid \tau) dz'}_{\text{Inflows due to separations}}$$
$$\underbrace{-p(\theta_t) \sum_{\tau' \in \mathcal{T}_t} \int \mathbb{1}^{UE}(h, z' \mid \tau, \tau'; t) \frac{v_t(z' \mid \tau')}{\mathbb{V}_t} g_t^v(z' \mid \tau') dz' g_t^u(h \mid \tau)}_{\text{Outflows due to U-E move}} \tag{D.20}$$
$$\underbrace{-\varrho g_t^u(h \mid \tau) + \varrho f_t(h \mid \tau)}_{\text{net retirement flows}}.$$

Specifically, for $\tau = 1$ this reduces to

$$\partial_t g_t^u(h \mid 1) = \underbrace{-\alpha_u \partial_h g_t^u(h \mid 1)}_{\text{Human capital depreciation}}$$
$$\underbrace{+\left(\kappa + \varrho \min\left\{\omega, \frac{1-u_t}{\iint g_t^e(z, h \mid 1)}\right\}\right) \int g_t^e(z', h \mid 1) dz'}_{\text{Inflows due to separations}} \tag{D.21}$$
$$\underbrace{-p(\theta_t) \sum_{\tau' \in \mathcal{T}_t} \int \mathbb{1}^{UE}(h, z' \mid 1, \tau'; t) \frac{v_t(z' \mid \tau')}{\mathbb{V}_t} g_t^v(z' \mid \tau') dz' g_t^u(h \mid 1)}_{\text{Outflows due to U-E move}}.$$

For $\tau = 0$ we get

$$\partial_t g_t^u(h \mid 0) = \underbrace{-\alpha_u \partial_h g_t^u(h \mid 0)}_{\text{Human capital depreciation}}$$
$$\underbrace{+\left(\kappa + \varrho \frac{\max\left\{1 - u_t - \omega \iint g_t^e(z, h \mid 1) dz dh, 0\right\}}{\iint g_t^e(z, h \mid 0) dz dh}\right) \int g_t^e(z', h \mid 0) dz'}_{\text{Inflows due to separations}} \tag{D.22}$$
$$\underbrace{-p(\theta_t) \sum_{\tau' \in \mathcal{T}_t} \int \mathbb{1}^{UE}(h, z' \mid 0, \tau'; t) \frac{v_t(z' \mid \tau')}{\mathbb{V}_t} g_t^v(z' \mid \tau') dz' g_t^u(h \mid 0)}_{\text{Outflows due to U-E move}}.$$

From the above expressions, it becomes clear that when looking at unemployment dynamics an exogenous increase in the skill supply of new workers is isomorphic to separation rates which depend on the technology vintage. In the special case where $\omega = 1$, the distributional dynamics above would be isomorphic to the version presented in the main text. Away from the special case, i.e. $\omega > 1$, the above expression acts to increase the overall supply of skilled workers but does so only slowly as the stock of workers experienced in the new technology increases.

Taking stock. Unemployment in the model in the main text can be viewed in a reduced form way as capturing a wide range of slow moving underlying mechanisms which increase the supply of workers experienced in the new technology as the transition unfolds. Furthermore, it captures that many such mechanisms are increasing the share of existing technology workers.

D.3 Alternative bargaining protocol

TO BE COMPLETED.

E Social Planner Solution

Preliminaries. This section extends and generalizes Fukui and Mukoyama (2025) to the current setting with heterogeneous human capital, learning and technology choice. Let Λ_t denote the social value of a match, Θ_t the social value of a solo entrepreneur, and Ω_t the social value of an unemployed. Let the operator $\mathcal{A}_{ee}$ denote all within match state changes that do not involve any worker flows and captured in Equation (5). Likewise, define the operator $\mathcal{A}_{vv}$ and $\mathcal{A}_{uu}$.

Derivations. Using standard methods we can apply integration by parts to derive the Lagrangian of the social planner problem as

$$
\begin{aligned}
\mathcal{L} = \max_{v,\mathbb{1}^{EE},\mathbb{1}^{UE},\mathbb{1}^{A},\mathbb{1}^{X}} & \int e^{-\rho t} \sum_{\tau\in\mathcal{T}_t} \int \pi(z,h \mid \tau) dG_t^e(z,h \mid \tau) dt \\
& + \int e^{-\rho t} \sum_{\tau\in\mathcal{T}_t} \int \left(\pi(z,0 \mid \tau) - c(v_t(z \mid \tau))\right) dG_t^v(z \mid \tau) dt \\
& + \int e^{-\rho t} \sum_{\tau\in\mathcal{T}_t} \int bh dG_t^u(h \mid \tau) dt \\
& + \int e^{-\rho t} \sum_{\tau\in\mathcal{T}_t} \int \mathcal{A}_{ee}\left[\Lambda_t\right](z,h \mid \tau) dG_t^e(z,h \mid \tau) dt \\
& + \int e^{-\rho t} \int \chi \mathbb{1}^{A}(z,h \mid \tau,\tau;t) \left(\mathbb{E}\left[\Lambda_t(z',h' \mid 1) \mid \tau,\tau\right] - \Lambda_t(z,h \mid 0)\right) dG_t^e(z,h \mid \tau) dt \\
& + \int e^{-\rho t} \sum_{\tau\in\mathcal{T}_t} \int \varkappa \mathbb{1}^{X}(z,h \mid \tau;t) \left(\Theta_t(z \mid \tau) + \Omega_t(h \mid \tau) - \Lambda_t(z,h \mid \tau)\right) dG_t^e(z,h \mid \tau) dt \\
& - \int e^{-\rho t} \sum_{\tau\in\mathcal{T}_t} \int \kappa \left(\Lambda_t(z,h \mid \tau) - \Theta_t(z \mid \tau) - \Omega_t(h \mid \tau)\right) dG_t^e(z,h \mid \tau) dt \\
& - \int e^{-\rho t} \sum_{\tau,\tau'\in\mathcal{T}_t} \iint \xi p(\theta_t) \left(\mathbb{1}^{EE}(z,h,z' \mid \tau,\tau';t) \frac{v_t(z' \mid \tau')}{\mathbb{V}_t} dG_t^v(z' \mid \tau')\right) \left(\Lambda_t(z,h \mid \tau) - \Theta_t(z \mid \tau)\right) dG_t^e(z,h \mid \tau) dt
\end{aligned}
$$

$$+\int e^{-\rho t}\sum_{\tau\in\mathcal{T}_t}\int\left(-\rho\Lambda_t(z,h\mid\tau)+\dot{\Lambda}_t(z,h\mid\tau)\right)dG_t^e(z,h\mid\tau)dt$$

$$+\int e^{-\rho t}\sum_{\tau\in\mathcal{T}_t}\int\mathcal{A}_{vv}\left[\Theta_t\right](z\mid\tau)dG_t^v(z\mid\tau)dt$$

$$+\int e^{-\rho t}\sum_{\tau,\tau'\in\mathcal{T}_t}\iint\xi p(\theta_t)\left(\mathbb{1}^{EE}(z',h',z\mid\tau';\tau,t)\left(\Lambda_t^H(z,h'\mid\tau,\tau')-\Theta_t(z\mid\tau)\right)dG_t^e(z',h'\mid\tau')\right)\frac{v_t(z\mid\tau)}{\mathbb{V}_t}dG_t^v(z\mid\tau)dt$$

$$+\int e^{-\rho t}\sum_{\tau,\tau'\in\mathcal{T}_t}\iint p(\theta_t)\left(\mathbb{1}^{UE}(h,z\mid\tau',\tau;t)\left(\Lambda_t^H(z,h'\mid\tau,\tau')-\Theta_t(z\mid\tau)-\Omega_t(h'\mid\tau')\right)dG_t^u(h'\mid\tau')\right)\frac{v_t(z\mid\tau)}{\mathbb{V}_t}dG_t^v(z\mid\tau)dt$$

$$+\int e^{-\rho t}\sum_{\tau\in\mathcal{T}_t}\int\left(-\rho\Theta_t(z\mid\tau)+\dot{\Theta}_t(z)\right)dG_t^v(z\mid\tau)dt$$

$$+\int e^{-\rho t}\sum_{\tau\in\mathcal{T}_t}\int\mathcal{A}_{uu}\left[\Omega_t\right](h\mid\tau)dG_t^u(h\mid\tau)dt$$

$$+\int e^{-\rho t}\sum_{\tau\in\mathcal{T}_t}\int\left(-\rho\Omega_t(h\mid\tau)+\dot{\Omega}_t(h\mid\tau)\right)dG_t^u(h\mid\tau)dt$$

$$+\int e^{-\rho t}\lambda_t\left(\theta_t\mathbb{X}_t-\mathbb{V}_t\right)dt$$

$$+\int e^{-\rho t}\mu_t\left(\sum_{\tau\in\mathcal{T}_t}\int dG_t^u(h\mid\tau)+\xi\sum_{\tau\in\mathcal{T}_t}\int dG_t^e(z,h\mid\tau)-\mathbb{X}_t\right)dt$$

$$+\int e^{-\rho t}\omega_t\left(\sum_{\tau\in\mathcal{T}_t}\int v_t(z\mid\tau)dG_t^v(z\mid\tau)-\mathbb{V}_t\right)dt$$

$$+\int e^{-\rho t}\phi_t\left(\sum_{\tau\in\mathcal{T}_t}\int dG_t^e(z,h\mid\tau)+\sum_{\tau\in\mathcal{T}_t}\int dG_t^u(h\mid\tau)-1\right)dt$$

$$+\int e^{-\rho t}\varphi_t\left(\sum_{\tau\in\mathcal{T}_t}\int dG_t^e(z,h\mid\tau)+\sum_{\tau\in\mathcal{T}_t}\int dG_t^v(z\mid\tau)-M\right)dt,$$

subject to

$$\Lambda_t^H(z,h\mid\tau,\tau')=\mathbb{1}^A(z,h\mid\tau,\tau',t)\,\mathbb{E}\left[\Lambda_t(z',h'\mid 1)\mid z,h,\tau,\tau'\right]+\left(1-\mathbb{1}^A(z,h\mid\tau,\tau',t)\right)\Lambda_t(z,h\mid 0).$$

Social value of a match. By standard Calculus of Variation arguments the social value functions must satisfy the following set of PDEs

$$\begin{aligned}\rho\Lambda_t(z,h\mid\tau)-\partial_t\Lambda_t(z,h\mid\tau)=\;&\pi(z,h\mid\tau)+\mathcal{A}_{ee}\left[\Lambda_t\right](z,h\mid\tau)-\kappa\left(\Lambda_t(z,h\mid\tau)-\Theta_t(z\mid\tau)-\Omega_t(h\mid\tau)\right)\\&+\xi p(\theta_t)\sum_{\tau'\in\mathcal{T}_t}\int\mathbb{1}^{EE}\left(z,h,z'\mid\tau,\tau';t\right)\Delta_t^{EE}(z,h,z'\mid\tau,\tau')\frac{v_t(z'\mid\tau')}{\mathbb{V}_t}dG_t^v(z'\mid\tau')\\&+\xi\mu_t+\phi_t+\varphi_t\end{aligned}\tag{E.23}$$

where

$$\Delta_t^{EE}(z,h,z' \mid \tau,\tau') = \left(\Lambda_t^H(z',h \mid \tau',\tau) - \Theta_t(z' \mid \tau'\right) - \left(\Lambda_t(z,h \mid \tau) - \Theta_t(z \mid \tau)\right)$$

and Λ_t^H defined as above, and subject to the boundary conditions for adoption and endogenous separation subject to the boundary conditions for adoption and endogenous separation

$$\Lambda_t(z,h|\tau) \geq \Lambda_t^A(z,h \mid \tau,\tau), \qquad \text{and} \qquad \Lambda_t(z,h|\tau) \geq \Theta_t(z|\tau) + \Omega_t(h|\tau),$$

where the value of adopting is given by $\Lambda_t^A(z,h \mid \tau_f,\tau_w) \equiv \mathbb{E}\left[\Lambda_t(\tilde{z},\tilde{h} \mid 1) \mid z,h,\tau_f,\tau_w\right]$, where $\tilde{z} = \delta^{1-\tau_f} z$ and $\tilde{h} = \delta^{1-\tau_w} h$ and δ follows the stochastic process defined in (4). Equation (E.23) differs from the competitive equilibrium in (5) only in terms of the last two lines.

Social value of solo entrepreneur. The value of the solo entrepreneur is

$$\begin{aligned}\rho\Theta_t(z \mid \tau) - \partial_t\Theta_t(z, \mid \tau) = {} & \pi(z,0 \mid \tau) - c(v_t(z \mid \tau)) + \mathcal{A}_{vv}\left[\Theta_t\right](z \mid \tau) \\ & + \xi p(\theta_t)\frac{v_t(z \mid \tau)}{\mathbb{V}_t}\sum_{\tau' \in \mathcal{T}_t}\iint \mathbb{1}^{EE}(z',h',z \mid \tau',\tau,t)\Delta_t^{EE}(z',h',z \mid \tau',\tau)dG_t^e(z',h' \mid \tau') \\ & + p(\theta_t)\frac{v_t(z \mid \tau)}{\mathbb{V}_t}\sum_{\tau' \in \mathcal{T}_t}\int \mathbb{1}^{UE}(h',z \mid \tau',\tau,t)\Delta_t^{UE}(h',z \mid \tau',\tau)dG_t^u(h' \mid \tau') \\ & + \omega_t v_t(z \mid \tau) + \varphi_t\end{aligned} \tag{E.24}$$

where

$$\Delta_t^{UE}(h,z' \mid \tau,\tau') = \Lambda_t^H(z',h \mid \tau',\tau) - \Theta_t(z' \mid \tau') - \Omega_t(h \mid \tau).$$

Equation (E.24) differs from the competitive equilibrium in (6) only in terms of the final line.

Value of unemployment. The value of unemployment given by

$$\begin{aligned}\rho\Omega_t(h \mid \tau) - \partial_t\Omega_t(h \mid \tau) = bh + \mathcal{A}_{uu}\left[\Omega_t\right](h \mid \tau) & \\ + p(\theta_t)\sum_{\tau' \in \mathcal{T}_t}\int \frac{v_t(z' \mid \tau')}{\mathbb{V}_t}\mathbb{1}^{UE}\left(h,z' \mid \tau,\tau' t\right)\Delta_t^{UE}(h,z' \mid \tau,\tau')g_t^v(z' \mid \tau')dz' & \\ + \mu_t + \phi_t &\end{aligned} \tag{E.25}$$

This PDE in (E.25) for the social value of unemployment again differs from the competitive equilibrium counterpart by their final two lines.

Optimal policies. Optimal vacancy posting satisfies

$$c'(v_t(z \mid \tau)) - \omega_t = \xi p(\theta_t) \frac{1}{\mathbb{V}_t} \sum_{\tau' \in \mathcal{T}_t} \iint \mathbb{1}^{EE} \left(z', h', z \mid \tau', \tau, t\right) \Delta_t^{EE} \left(z', h', z \mid \tau', \tau, \right) dG_t^e(z', h' \mid \tau')$$
$$+ p(\theta_t) \frac{1}{\mathbb{V}_t} \sum_{\tau' \in \mathcal{T}_t} \int \mathbb{1}^{UE} \left(h', z \mid \tau', \tau, t\right) \Delta_t^{UE} \left(h', z \mid \tau', \tau, \right) dG_t^u(h' \mid \tau').$$

Optimal hiring and adoption policy then simply is given by

$$\mathbb{1}^{EE}(z, h, z' \mid \tau, \tau'; t) = \begin{cases} 1 & \text{if } \Delta^{EE}(z, h, z' \mid \tau, \tau') \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{1}^{UE}(h, z' \mid \tau, \tau'; t) = \begin{cases} 1 & \text{if } \Delta^{UE}(h, z' \mid \tau, \tau') \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{1}^{A}(z, h' \mid \tau, \tau'; t) = \begin{cases} 1 & \text{if } \Lambda_t^A(z, h \mid \tau, \tau') \geq \Lambda_t(z, h' \mid 0) \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbb{1}^{X}(z, h \mid \tau; t) = \begin{cases} 1 & \text{if } \Lambda_t(z, h \mid \tau) \leq \Theta_t(z \mid \tau) + \Omega_t(h \mid \tau) \\ 0 & \text{otherwise.} \end{cases}$$

Congestion Lagrange multipliers satisfy

$$-\omega_t \mathbb{V}_t = \xi p(\theta) \left(1 - \frac{p'(\theta_t)\theta_t}{p(\theta_t)}\right) \sum_{\tau, \tau' \in \mathcal{T}_t} \int \frac{v_t(z \mid \tau)}{\mathbb{V}_t} \Delta_t^{EE} \left(z', h', z \mid \tau', \tau\right)^+ dG_t^e(z', h' \mid \tau') dG_t^v(z \mid \tau)$$
$$+ p(\theta) \left(1 - \frac{p'(\theta_t)\theta_t}{p(\theta_t)}\right) \sum_{\tau, \tau' \in \mathcal{T}_t} \int \frac{v_t(z \mid \tau)}{\mathbb{V}_t} \Delta_t^{UE} \left(h', z \mid \tau', \tau\right)^+ dG_t^u(h' \mid \tau') dG_t^v(z \mid \tau)$$

The worker search congestion is given by

$$-\mu_t \mathbb{X}_t = \xi p(\theta_t) \left(\frac{p'(\theta_t)\theta_t}{p(\theta_t)}\right) \sum_{\tau, \tau' \in \mathcal{T}_t} \int \frac{v_t(z \mid \tau)}{\mathbb{V}_t} \Delta_t^{EE}(z', h', z \mid \tau', \tau)^+ dG_t^e(z', h' \mid \tau') dG_t^v(z \mid \tau)$$
$$+ p(\theta_t) \left(\frac{p'(\theta_t)\theta_t}{p(\theta_t)}\right) \sum_{\tau, \tau' \in \mathcal{T}_t} \int \frac{v_t(z \mid \tau)}{\mathbb{V}_t} \Delta_t^{UE}(h', z \mid \tau', \tau)^+ dG_t^u(h' \mid \tau') dG_t^v(z \mid \tau)$$

Note that these congestion effects relate to elasticity of the matching function. This mirrors Hosios (1990).

Optimal taxes. These can be backed out by comparing the social planner and competitive equilibrium value functions.

F Computational Appendix

F.1 Stationary equilibrium

This section describes the computational algorithm for solving the stationary equilibrium of the model with $\mathcal{T} = \{0\}$. The algorithm employs a nested fixed-point iteration structure, solving the coupled system of Hamilton-Jacobi-Bellman (HJB) equations and Kolmogorov Forward Equations (KFE) in steady-state. There are a set of distinct numerical challenges that the algorithm addresses: (i) the entire distribution enter as a state, (ii) the transition matrices are generically dense, (iii) the KFE consists of coupled non-linear system with mass-conservation constraints, and (iv) productivity and human capital distributions are an endogenous outcome.

Algorithm.

Step 0: *Set-up.* Discretize grids for the state variables. Let $h \in [\underline{h}, \bar{h}]$ and $z \in [\underline{z}, \bar{z}]$. Transform into $\tilde{h} = \log(h) - \mathbb{E}\log(h)$ and $\tilde{z} = \log(z) - \mathbb{E}\log(z)$, where $\mathbb{E}$ is computed using the endogenous ergodic distributions. This transformation allows to use a grid centered around 0. Discretize the computational space using $N_h \times N_z$ grid points with uniform spacing Δh and Δz.

Step 1: *Initialize.* Given the discretized grid, make an initial guess for the vacancy posting policy $v_{ss}^{(0)}$, distribution vectors discretized on a histogram $\boldsymbol{g}_{ss}^{(0)} = \{g_{ss}^{e,(0)}, g_{ss}^{v,(0)}, g_{ss}^{u,(0)}\}$, and approximate averages $\overline{h}^{(0)}, \overline{z}^{(0)}, Y^{(0)}$. Collect $\boldsymbol{y}^{(0)} = \{v_{ss}^{(0)}, \boldsymbol{g}_{ss}^{(0)}\}$. Choose convergence tolerance ε.

Step 2: *Update.* For each iteration $n \geq 1$, update aggregate variables $\mathbb{V}_{ss}, \mathbb{X}_{ss}, \theta_{ss}, u_{ss}$.

Step 3: *Solve HJB.* Given aggregates and $\boldsymbol{y}^{(n)}$ and normalized grids update value functions S_{ss} and V_{ss}. Apply grid normalization using $\overline{h}^{(n)}$ and $\overline{z}^{(n)}$. Normalize value functions by aggregate output $Y^{(n)}$.

3.1 Discretize HJB using finite differences.

3.2 Use implicit-explicit (IMEX) scheme with adaptive time-stepping, treating the vacancy posting gain as explicit, to handle dense transition matrices.

3.3 Use sparse automatic differentiation to assemble Jacobians.

3.4 Use GMRES with incomplete LU preconditioning.

3.5 Handle optimal stopping times using penalty functions.

3.6 Iterate until HJB residual is sufficiently small or maximum iterations reached. Set $\text{flag}_{HJB} = \text{true}$ if converged.

Step 4: *Solve KFE.* To solve the coupled non-linear system, iterate on:

4.1 Update entrepreneur distributions. Use augmented system enforcing total mass of M. Use GMRES with incomplete LU preconditioning.

4.2 Update worker distributions, imposing one worker-one entrepreneur constraint using updated entrepreneur distribution.

4.3 Check convergences. If converged return flag_{KFE} = true. Continue to 5.

4.4 Enforce total mass of workers to equal 1. Perform dampened update of distribution. Iterate until convergence or maximum iterations reached.

Step 5: *Convergence check.* Based on HJB and KFE solutions compute new predicted $\boldsymbol{y}^{pred}$. Check convergence. If $\|\mathbf{y}^{(n)} - \mathbf{y}^{\text{pred}}\| < \varepsilon$ and $\text{flag}_{HJB} \wedge \text{flag}_{KFE}$, terminate.

Step 6: *Update.* Compute $\boldsymbol{y}^{(n+1)}$ using Type I Anderson acceleration. Use Tiakonov Regularization if ill-conditioned. Use mixing and safeguarding. Re-enforce mass constraints on distribution. Update approximate averages $\overline{h}^{(0)}$ and $\overline{z}^{(0)}$. Set $n \leftarrow n + 1$.

F.2 Transition

I solve for the transition path after the unexpected introduction of a new technology, i.e. the expansion of the set $\mathcal{T}_t$. A key challenge in the transition is to enforce the mass constraints and coupling between worker and entrepreneur distributions at every step along the transition. I employ a Predictor-Corrector method.

Algorithm.

Step 0: Compute the stationary equilibrium before, i.e. featuring $\mathcal{T}_t = \{0\}$, and after the introduction of the new technology, i.e. featuring $\mathcal{T}_t = \{0, 1\}$.[50]

Step 1: Truncate the horizon at a sufficiently large T. Guess a sequence for labor market tightness $\{\theta_t^{(0)}\}_{t=0}^T$ and distributions $\{\boldsymbol{g}_t^{(0)}\}$.

Step 2: For iteration $n \geq 1$, update aggregate variables $\{\mathbb{V}_t, \mathbb{X}_t, u_t\}_{t=0}^T$.

Step 3: Solve backwards for $\{\boldsymbol{x}_t\}_{t=0}^T \equiv \{S_t, V_t\}_{t=0}^T$ using the HJB. The procedure is similar to the stationary solution with a minor modification. The HJB at each time step corresponds to solving a non-linear system $\mathcal{F}(\boldsymbol{x}_t, \boldsymbol{x}_t^*, \boldsymbol{x}_{t+\Delta}; \Theta) = 0$. Since computing the entire Jacobian of this system is computationally costly due to its non-sparse nature, I solve this for each time step using a predictor-corrector approach. That is, set $\hat{\boldsymbol{x}}_t^{0,n} = \boldsymbol{x}_{t+\Delta}^{(n)}$ and iterate on:

3.1 Solve for $\hat{\boldsymbol{x}}_t$ in $\mathcal{F}(\hat{\boldsymbol{x}}_t, \hat{\boldsymbol{x}}_t^{r,n}, \boldsymbol{x}_{t+\Delta}; \Theta) = 0$. Given the IMEX structure this is simple and fast.

3.2 Check if $\|\hat{\boldsymbol{x}}_t - \hat{\boldsymbol{x}}_t^{r,n}\| < \varepsilon$. If converged, set $\boldsymbol{x}_t^{(n)} = \hat{\boldsymbol{x}}_t$.

[50] At the end, it is important to verify that the economy indeed converges to this new equilibrium.

3.3 Set $r \leftarrow r + 1$.

Step 4: Solve forwards for $\{\boldsymbol{g}_t^{(n+1/2)}\}_{t=0}^T$. Enforce mass constraints and improve accuracy by using a predictor-corrector approach similar to the above.

Step 5: Based on $\boldsymbol{x}_t^{(n)}$ and new updates $\{\boldsymbol{g}_t^{(n+1/2)}\}_{t=0}^T$, compute new predicted $\{\theta_t^{(n+1/2)}\}_{t=0}^T$. Check convergence in all three sequences. If converged, terminate.

Step 6: Compute update to $\{\boldsymbol{g}_t^{(n+1)}\}_{t=0}^T$ and $\{\theta_t^{(n+1)}\}_{t=0}^T$ using Type I Anderson acceleration. Use Tiakonov Regularization if ill-conditioned. Use mixing and safeguarding. Re-enforce mass constraints on distribution.

G Analytical Model

This section illustrates the trade-off between production complementarities, dynamic learning and match duration in an analytical model while capturing the main elements of the quantitative model. For most of the section we focus on the stylized case where there are two entrepreneurs and a single worker. This setting allows us to think through the key forces governing the optimal allocation of the worker between entrepreneurs as well as the optimal match duration that a social planner chooses. Towards the end of the section, we introduce a third entrepreneur C to understand the externalities and inefficiencies in the competitive equilibrium with Bertrand competition and learning.

There are two entrepreneurs A and B and a single worker. Time is continuous. Productivity is denoted by $z_A(t)$ and $z_B(t)$. The worker's human capital is denoted by $h(t)$. Suppose that $z_B(0) < h(0) < z_A(0)$, and $h(0) > 1$. How should the worker be allocated between entrepreneur A and B?

Production. Production features complementarities between z and h parameterized by χ. Specifically, output for firm $i \in \{A, B\}$ is given by

$$y_i(t) = \begin{cases} z_i(t) + h(t) + \chi z_i(t) h(t) & \text{if matched with worker} \\ z_i(t) & \text{otherwise.} \end{cases}$$

Learning dynamics. Learning happens from the more to the less knowledgeable agent and is increasing in the difference of the knowledge levels. Specifically, learning takes the form for the entrepreneur

$$\dot{z}_i(t) = \begin{cases} \alpha_1 \left[h(t) - z_i(t)\right]^+ & \text{if matched with worker} \\ 0 & \text{otherwise.} \end{cases}$$

For the worker matched with entrepreneur i human capital evolves as

$$\dot{h}(t) = \alpha_1 \left[z_i(t) - h(t)\right]^+ .$$

If a match were to last forever, the learning dynamics imply that the present value of productivity and human capital is a simple weighted average of the initial knowledge and the knowledge of the more knowledgeable peer. Specifically,

$$\int_0^{+\infty} e^{-rt} h(t) dt = \frac{r}{r+\alpha_1}\frac{h(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{z_A(0)}{r} \qquad (h \text{ learns from } A)$$

$$\int_0^{+\infty} e^{-rt} z_B(t) dt = \frac{r}{r+\alpha_1}\frac{z_B(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{h(0)}{r} \qquad (B \text{ learns from } h)$$

$$\int_0^{+\infty} e^{-rt} z_A(t) dt = \frac{z_A(0)}{r} \qquad (A \text{ does not learn})$$

Once and for all allocation. First, we consider the optimal time zero allocation of the worker assuming that it is permanent. The gain from allocating the worker to A comes from two sources: (i) a production misallocation effect modulated through the complementarity parameter χ as the production function is supermodular and (ii) a learning effect of h learning from the more productive firm. The gain from allocating the worker to B is a pure learning effect, i.e. from raising the productivity of the laggard. The following Lemma formalizes this.

Lemma 5. *The net present value of output from allocating to entrepreneur A and B, respectively, are*

$$V_A = \frac{z_A(0)}{r} + \frac{z_B(0)}{r} + \left(\frac{r}{r+\alpha_1}\frac{h(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{z_A(0)}{r}\right) + \chi\left(\frac{r}{r+\alpha_1}\frac{z_A(0)h(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{z_A(0)^2}{r}\right)$$

$$V_B = \frac{z_A(0)}{r} + \left(\frac{r}{r+\alpha_1}\frac{z_B(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{h(0)}{r}\right) + \frac{h(0)}{r} + \chi\left(\frac{r}{r+\alpha_1}\frac{z_B(0)h(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{h(0)^2}{r}\right).$$

It is optimal to allocate the worker to the less productivity entrepreneur B if and only if

$$\underbrace{\frac{\alpha_1}{r+\alpha_1}\left(\frac{h(0)}{r} - \frac{z_B(0)}{r}\right)}_{\text{gain } B \text{ learning}} - \underbrace{\frac{\alpha_1}{r+\alpha_1}\left(\frac{z_A(0)}{r} - \frac{h(0)}{r}\right)}_{\text{gain } h \text{ learning}} > \underbrace{\chi\frac{(z_A(0) - z_B(0))\, h(0)}{r+\alpha_1}}_{\text{static misallocation}} + \underbrace{\chi\frac{\alpha_1}{r+\alpha_1}\left(\frac{z_A(0)^2}{r} - \frac{h(0)^2}{r}\right)}_{\text{dynamic misallocation}} \tag{G.26}$$

The left hand side in G.26 motivates worker mobility and diffusion [better explanation of intuition needed]. If the learning gains for B from h exceed the learning gains of h from A, then there is a motive for allocating the worker to B. The right hand side denotes the sum of static and dynamic production misallocation costs. For $\chi > 0$ there are production complementarities between z and h. The first expression captures the losses from misallocating at initial knowledge levels. The second dynamic expression captures that these losses are compounded by the fact that the worker becomes more knowledgeable over time at A.

It is instructive to consider limiting cases. If there is no learning or complementarities in production are large, i.e. $\alpha_1/\chi \to 0$, then the condition can never be satisfied given $z_A(0) > \max\{z_B(0), h(0)\}$ as only the misallocation terms remain. When complementarities in production are shut off, i.e. $\chi \to 0$, then the above condition reduces to

$$\left(\frac{h(0)}{r} - \frac{z_B(0)}{r}\right) - \left(\frac{z_A(0)}{r} - \frac{h(0)}{r}\right) > 0.$$

Intuitively, with no consideration for production complementarities the social planner wants to allocate the worker to maximize total knowledge and therefore exploit all potential learning gains. Away from this limit case the social planner trades off production value and learning gains.

Optimal match duration. The previous equation motivates the following thought exercise. What is the optimal duration T that the worker should stay at firm A before moving to B? The social planner maximizes

$$\max_T \int_0^T e^{-rt}\left(z_A(0) + h(t) + z_B(0)\right) dt + \chi \int_0^T e^{-rt} z_A(0) h(t) dt + e^{-rt} V_B(h(T)),$$

where the continuation value of allocating the worker with human capital to firm B is given by

$$V_B(h) = \frac{z_A(0)}{r} + \frac{h}{r} + \left(\frac{r}{r+\alpha_1}\frac{z_B(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{h}{r}\right) + \chi h \left(\frac{r}{r+\alpha_1}\frac{z_B(0)}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{h}{r}\right).$$

Its derivative is

$$V_B'(h) = \frac{1}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{1}{r} + \chi\frac{z_B(0)}{r+\alpha_1} + 2\chi\frac{\alpha_1}{r+\alpha_1}\frac{h}{r}.$$

The first order condition with respect to T is given by

$$\underbrace{z_A(0) + h(T) + z_B(0) + \chi z_A(0) h(T) +}_{\text{(1) flow return of delay}} \underbrace{V_B'(h(T)) h'(T)}_{\text{(2) increased learning gains for } B} - \underbrace{r V_B(h(T))}_{\text{(3) cost of delay}} = 0.$$

Noting that $h'(T) = \alpha_1 (z_A(0) - h(T))$. The above boils down to finding the optimal threshold level of h^* that satisfies the above equation. Substituting in for the functional forms we obtain the quadratic equation

$$\begin{aligned}
0 = -\frac{\alpha_1}{r+\alpha_1}\left(h^* - z_B(0)\right) &+ \chi\frac{r}{r+\alpha_1}\left(z_A(0) - z_B(0)\right) h^* \\
&+ \chi\frac{\alpha_1}{r+\alpha_1}\left(z_A(0) - h^*\right) h^* \\
&+ \alpha_1\left(\frac{1}{r} + \frac{\alpha_1}{r+\alpha_1}\frac{1}{r} + \chi\frac{z_B(0)}{r+\alpha_1} + 2\chi\frac{\alpha_1}{r+\alpha_1}\frac{h^*}{r}\right)\left(z_A(0) - h^*\right).
\end{aligned}$$

A sufficient condition for finite T to exist is $z_A(0) < \alpha_1/(\chi r)$. Proof: Plugin in for $h = z_A(0)$. Likewise, we can plug in for $h = z_B(0)$ and show that the $LHS > 0$. Then, can use the intermediate value theorem.

Once we have found the optimal h^* we can back out the optimal match duration

$$T(h^*) = \frac{1}{\alpha_1} \log\left(\frac{z_A(0) - h(0)}{z_A(0) - h^*}\right).$$

Discussion. Let us interpret these motives. Delaying the reallocation comes with two benefits. First, the worker is allocated to the more productive entrepreneur A. In the presence of production complementarities this yields large contemporaneous aggregate output. Second, insofar as that entrepreneur A is more productive than the worker there are additional gains from the worker remaining at A as this allows for the human capital to increase. Crucially, this gives another motive for match duration in the model. On the other hand, the cost of delay stems from the missed learning opportunities by entrepreneur B.

Competitive equilibrium. I now illustrate the sources of inefficient worker allocations in a setting of the bargaining protocol by Postel-Vinay and Robin (2002). Entrepreneurs Bertrand compete on wages for the worker such that the worker moves to the entrepreneur where they have the greatest marginal surplus. As a result, the bargaining protocol is bilaterally efficient and all learning gains between A and B are fully internalized. Therefore, to be study the role of knowledge spillovers we introduce a third entrepreneur C. Anticipating the results below, the primary inefficiency comes from entrepreneurs A and B not internalizing how their hiring decisions affect the endogenous creation of knowledge which may spill over to a prospective poachee C.

For illustrative purposes, let $z_B(0) \leq z_C(0) < h(0) < z_A(0)$. B and C have the potential to learn from h and h has the potential to learn from A. To isolate the key inefficiency we assume the following deterministic labor market setting. At time 0 the worker is allocated to entrepreneur B and there is an opportunity of A to poach the worker. After a time interval T there is a meeting between the worker and the entrepreneur C who gets the opportunity to poach the worker. The following event tree summarizes the event tree.

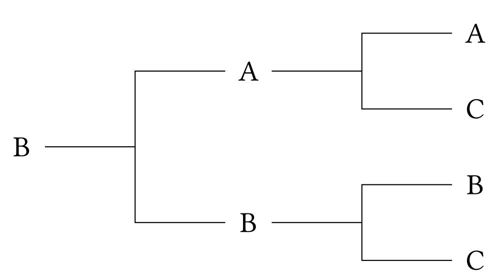


Figure G.9: Event sequences in labor market.

Before turning to the inefficiencies, it is instructive to first describe the allocations in the competitive equilibrium the sequential auction bargaining protocol by Postel-Vinay and Robin (2002). If the worker is at A inititially and after a time period T moves to C then based on the bargaining protocol the entrepreneur and worker value are, respectively,

$$J_0(A \mid h) = \int_0^T e^{-rt}\left(z_A(0) + h(t) + \chi z_A(0)h(t) - w_t\right) + \int_T^{+\infty} e^{-rt} z_A(0)dt$$
$$W_0(h \mid A) = \int_0^T e^{-rt} w_t dt + W_T(h \mid C)$$

where the sequential auction protocol implies that the worker extracts the value of the past employer A from the new employer C. That is, the maximum value A would be willing to pay to employ the worker $W_T(h \mid C) = \int_T^{+\infty} e^{-rt}\left(z_A(0) + h(t) + \chi z_A(0)h(t) - z_A(0)\right)dt$. By the Bertrand assumption this is the wage that C offers if poaching was optimal. Combined, the total private surplus is given by the following expression that is independent of whether C poaches the worker

$$S_0(A, h) = J_0(A \mid h) + W_0(h \mid A) = \int_0^{+\infty} e^{-rt}\left(z_A(0) + h(t) + \chi z_A(0)h(t) - z_B(0)\right)dt.$$

Likewise, the total private surplus of the worker remaining at B is

$$S_0(B, h) = \int_0^{+\infty} e^{-rt}\left(z_B(t) + h(0) + \chi z_B(t)h(0) - z_B(0)\right)dt.$$

Again by Bertrand competition entrepreneurs are able to pay up to their full surplus to employ the worker. As utility is perfectly transferable the worker will therefore move to the entrepreneur with the greatest surplus. It follows that in the competitive equilibrium the worker remains at entrepreneur B if and only if

$$\int_0^{+\infty} e^{-rt}\left(z_B(t) + h(0) + \chi z_B(t)h(0) - z_B(0)\right)dt > \int_0^{+\infty} e^{-rt}\left(z_A(0) + h(t) + \chi z_A(0)h(t) - z_B(0)\right)dt.$$

We note that this is identical to the condition discussed in the bilateral case and all results carry over. It follows that in absence of a third entrepreneur C the competitive equilibrium is perfectly efficient.

Next let us turn to the social planner solution

Lemma 6. *The social value of moving the worker to A is given by*

$$W_A^{SP} = \underbrace{\int_0^{+\infty} e^{-rt}\left(z_A(0) + h(t) + \chi z_A(0)h(t)\right)dt + \int_0^{\infty} e^{-rt} z_B(0)dt + \int_0^{+\infty} e^{-rt} z_C(0)dt}_{\text{private value}}$$
$$+ \underbrace{\int_T^{+\infty} e^{-rt}\left(z_C(t \mid A) + h(T \mid A) + \chi z_C(t \mid A)h(T \mid A) - z_C(0)\right)dt}_{\text{added surplus from moving worker to } C \text{ (including learning)}}$$
$$- \underbrace{\int_T^{+\infty} e^{-rt}\left(h(t) + \chi z_A(0)h(t)\right)dt}_{\text{social cost of moving worker away from } A}$$

Some discussion here. E.g. that sum of last two lines is positive. That matters how much residual learning. Also on χ such that large χ accentuates the production misallocation.

Likewise,

$$W_B^{SP} = \underbrace{\int_0^{+\infty} e^{-rt}\left(z_A(0)dt + z_B(t \mid B) + h(0) + \chi z_B(t \mid B)h(0)\right)dt + \int_0^{+\infty} e^{-rt} z_C(0)dt}_{\text{private value}}$$
$$+ \underbrace{\int_T^{+\infty} e^{-rt}\left(z_C(t \mid B) + h(0) + \chi z_C(t \mid B)h(0) - z_C(0)\right)dt}_{\text{added surplus from moving worker to } C \text{ (including learning)}}$$
$$- \underbrace{\int_T^{+\infty} e^{-rt}\left(z_B(t \mid B) + h(0) + \chi z_B(t \mid B)h(0) - z_B(T \mid B)\right)dt}_{\text{social cost of moving worker away from } B}$$

It then follows that the wedge between the social planner and the competitive equilibrium are

$$e^{rT}\left(W_B^{SP}-W_A^{SP}-(W_B^{CE}-W_A^{CE})\right)=\underbrace{-\int_T^{+\infty}e^{-rt}\left(z_C(t\mid A)-z_C(t)\right)dt}_{\text{greater learning for } C \text{ if first at } A}$$
$$\underbrace{-\chi\int_T^{+\infty}e^{-rt}\left(z_C(t\mid A)h(T\mid A)-z_C(t\mid B)h(0)\right)dt}_{\text{misallocation due to greater learning if first at } A}$$
$$\underbrace{-\int_T^{+\infty}e^{-rt}\left(z_B(t\mid B)-z_B(T\mid B)\right)dt}_{\text{foregone learning at } B \text{ if move to } C}$$
$$\underbrace{+\int_T^{+\infty}e^{-rt}\left(h(t\mid A)-h(T\mid A)\right)dt}_{\text{foregone learning of } h \text{ if move to } C}$$
$$\underbrace{+\chi\int_T^{+\infty}e^{-rt}\left(z_A(0)h(t\mid A)-z_B(t\mid B)h(0)\right)dt}_{\text{misallocation due to greater productivities at } A}$$

Combining terms

$$e^{rT}\left(W_B^{SP}-W_A^{SP}-(W_B^{CE}-W_A^{CE})\right)=\underbrace{-\frac{\alpha_1}{r+\alpha_1}\left[\left(\frac{h(T\mid A)}{r}-\frac{h(0)}{r}\right)-\left(\left(\frac{z_A(0)}{r}-\frac{h(T\mid A)}{r}\right)-\left(\frac{h(0)}{r}-\frac{z_B(T)}{r}\right)\right)\right]}_{(1)}$$
$$\underbrace{+\chi\frac{r}{r+\alpha_1}\frac{(z_A(0)-z_C(0))\,h(T\mid A)-(z_B(T)-z_C(0))\,h(0)}{r}}_{(2)}$$
$$\underbrace{+\chi\frac{\alpha_1}{r+\alpha_1}\left(\frac{z_A(0)^2-h(T\mid A)^2}{r}\right)}_{(3)}$$

where (1) denotes the net learning gains and (2) and (3) the net misallocation.

Comment. TO BE COMPLETED.