On-the-job Learning and Automation in Career Paths

Meghana Gaur *

[Updated often, click here for the latest version]

November 2, 2026

Abstract

Young workers quickly transition to higher-paying and less routine occupations within the first decade of their careers. This paper studies how on-the-job learning shapes the impacts of automation shocks within and across career paths. I develop a dynamic general equilibrium model of the labor market in which young workers sort into occupations that differ in their learning opportunities. Combining the structure of the occupational choice model with large-scale résumé data, I recover a flexible on-the-job learning technology that rationalizes observed patterns of early-career occupational mobility. The calibrated model shows how an automation shock to one occupation can induce greater automation further up the occupational career ladder. For a broad-based automation shock, adjustments in workers' skill accumulation decisions provide an important margin of long-run labor-supply flexibility, substantially lowering the pass-through of automation exposure to relative wages.

1 Introduction

A growing concern accompanying recent advances in generative artificial intelligence (AI) is that early-career roles, which may be important for skill development, are particularly vulnerable to automation (Afrouzi et al. (2026), Ide (2026)). For example, working as a financial analyst or software programmer might equip young workers with the skills required for working in financial or software management positions later in their careers. While the former occupations are considered to be exposed to AI-driven automation, the latter are typically thought to be relatively less exposed, reflecting their greater emphasis on interpersonal tasks, higher-stakes decision-making, and judgment. As these particularly salient examples illustrate, measuring an individual's exposure to automation at a point in time does not tell us their future exposure to automation or the impacts of automation on their ultimate career trajectory.

Understanding the labor market impacts of new automation technologies that differentially affect labor demand for occupations within and across career paths requires a quantitative framework that accounts for (i) how experience in one occupation contributes to the accumulation of skills in another, (ii) the dynamic nature of occupational choice, and (iii) the economic incentives governing automation decisions. However, existing state-of-the-art quantitative frameworks for studying the distributional impacts of automation tend

*I am grateful to John Grigsby, Richard Rogerson, and Ayşegül Sahin for their guidance and support on this project. I would also like to thank participants from the macro and trade student workshops at Princeton for their thoughtful comments and gratefully acknowledge financial support from the Stripe Economics of AI Fellowship and National Science Foundation's Graduate Research Fellowship Program.

to either abstract from on-the-job skill accumulation or represent skills or occupations at a level of aggregation that may obscure important features of modern career paths.

This paper seeks to make progress on these issues and proceeds in four steps. First, I contribute descriptive evidence on the nature of early career occupational transitions, drawing on large-scale résumé data from LinkedIn job profiles provided by Revelio Labs. Second, I build a dynamic general equilibrium model of the labor market that makes central the role of initial occupational choice in human capital accumulation. Within this environment, I characterize how on-the-job learning between occupations shapes the long-run response of occupation-level labor supply to changes in relative skill prices and can even lead to complementarities in occupation-level supply. Third, I use the structure of the occupational choice model to recover a reduced-form learning technology that rationalizes patterns of occupational mobility in the job profiles data. Fourth, I use the model to quantitatively evaluate the potential labor market impacts of generative AI-driven automation shocks within and across career paths.

In the first part of the paper, I use data on LinkedIn job profiles, provided by Revelio Labs, to document descriptive patterns of early-career occupational mobility. I show that recent cohorts of U.S.-based LinkedIn users quickly transition to higher-paying occupations with greater non-routine interpersonal and analytical task content during the first decade of their careers. Changes in task content slow significantly after the first decade of experience and largely reflect a life-cycle shift toward managerial work, in line with recent work by Deming (2025). I also show that, among users with a bachelor's degree, a common measure of automation exposure to generative AI falls with labor market experience.[1] The scale of the job profiles data allows me to provide granular examples of the typical transitions that contribute to these findings (e.g., financial analyst to financial manager). These features of professional career paths motivate this paper's focus on understanding the impacts of uneven automation shocks to occupations within and across career paths.

In the second part of the paper, I develop a dynamic model with on-the-job learning and automation, which builds on existing dynamic discrete choice models of occupational choice, such as those studied in Althoff and Reichardt (2026) and Dvorkin and Monge-Naranjo (2019). In the model, young workers enter the labor market with heterogeneous skills for different occupations. The key ingredient is a directed, bilateral on-the-job learning technology that allows workers to accumulate skills for occupations other than the one in which they are currently employed. Within this framework, I show how on-the-job learning can generate positive long-run cross-wage elasticities of labor supply, whereby an increase in the wage of one occupation *increases* labor supply to another occupation.

Two distinct mechanisms contribute to the possibility of positive cross-wage labor supply elasticities. First, forward-looking workers may respond to an increase in the skill price for financial managers by pursuing work as a financial analyst if it allows them to accumulate skills valuable for future managerial employment. Second, when an increase in the wage of financial analysts attracts young workers toward analyst work, it also increases the number of future workers skilled in managerial work. This second mechanism operates even if workers do not anticipate or value the future returns to the skills that they accumulate.

Positive cross-wage labor supply elasticities, which I refer to as *complementarities in labor supply*, are necessarily absent from static models without on-the-job learning. In static models, higher wages in one occupation lower labor supplied to others, all else equal. Complementarities in labor supply are an inherently dynamic concept as they depend on the forward-looking skill accumulation decisions of young workers and the dynamic

[1] Importantly, I document these patterns for cohorts that entered the labor market before the emergence of generative AI.

impacts of their occupational choices on the skill distribution of experienced workers. While such complementarities are likely to be present in existing models of dynamic occupational choice that are calibrated to match empirical patterns of life-cycle mobility (e.g., Althoff and Reichardt (2026) and Traiberman (2019)), I show that they have novel implications for patterns of automation across occupations in long-run equilibrium. In particular, an exogenous automation shock to one occupation will tend to increase automation in the occupations that are labor-supply complements to it, all else equal.

In the third part of the paper, I use the structure of the occupational choice model to recover a flexible on-the-job learning technology from occupational transition networks that I construct using the job profiles data. The large size of the data allows me to estimate long-run occupational transition networks separately for narrowly defined groups of workers and across a large number of occupations. In general, the amount of learning that I infer between two occupations strongly correlates with, but is only partially explained by, observable measures of skill, knowledge, and task similarity and industry overlap. These associations suggest that the accumulation of domain-specific human capital is an important feature of on-the-job learning in career paths. Importantly, I also find that the inferred learning technology is directional: learning tends to be larger for moves toward occupations with greater reasoning and decision-making content. This latter finding supports the notion of occupational career ladders.

The calibrated model confirms the relevance of positive cross-wage labor supply elasticities. In the model, complementarities in labor supply appear in (directed) pairs of occupations that constitute intuitive career paths. For example, in the model, labor supply in computer and information systems managerial work is *increasing* in the skill price for software developers, all else equal. More generally, I show that accounting for endogenous skill accumulation decisions substantially increases the overall magnitude of labor supply elasticities, relative to a static model of occupational choice without skill accumulation that is calibrated to match the same cross-sectional patterns of occupational sorting and wages, but ignores occupational histories. This simple comparison illustrates how endogenous skill accumulation decisions amplify the long-run responsiveness of labor supply to changes in relative skill prices.

Finally, in the last part of the paper, I use the model to evaluate the potential labor market impacts of generative AI-driven improvements in occupation-level automation technologies. I first consider an isolated automation shock directed at a single occupation: software developers. I focus on software developers for two reasons. First, it is an occupation that is commonly considered to be highly exposed to AI-driven automation. Second, I show that it is also an important entry point into an occupation with lower automation exposure: computer and information systems (CIS) managers. These patterns of occupational mobility lead the calibrated model to infer that the two occupations are (long-run) labor supply complements.

In the long run, an automation shock to software developer tasks thus induces additional automation of the tasks performed by CIS managers, following the simple intuition outlined above. However, the quantitative model shows that this automation response unfolds gradually. The increase in the skill price for CIS managers is delayed as it arises from the reduced entry of young workers into the software development career path. In the short run, however, labor supply reallocation pushes *down* the skill price for CIS managers as incumbent workers reallocate from software development work toward software managerial work as these workers face relatively similar earnings prospects across these two occupations, relative to other alternatives. Thus, the model clarifies that two occupations can look like labor-supply substitutes in the short-run and labor-supply complements in the long-run.

These gradual adjustments in skill prices have important distributional implications. I find that incumbent workers with a strong comparative advantage in software development work that entered the labor market before the automation shock are worse off. In contrast, future cohorts of workers with the same skill profile gain from the shock, as they are cushioned by the equilibrium adjustments in wages resulting from the reallocation of other workers away from the software development career path.

After examining the consequences of an isolated automation shock to a single occupation within an existing career ladder, I use the calibrated model to study the potential impacts of a broad-based automation shock. To conduct this experiment, I rely on existing occupation-level measures of AI-driven automation exposure to discipline the pattern of automation shocks that I feed into the model. I find that accounting for endogenous skill accumulation decisions substantially dampens the pass-through of automation exposure to skill prices, reducing the long-run dispersion in occupational skill price by 43% relative to the upon-impact response. This dampening largely redistributes the long-run gains of automation among workers with different types of skills. For example, workers with an initial comparative advantage in the occupations that expand in response to AI-driven automation, such as healthcare and service occupations, gain the most early on in the transition, while the reverse is true for workers with an initial comparative advantage in the cognitive occupations that are most exposed to automation.

Related literature. This paper relates to four strands of literature. Most importantly, it contributes to the literature that documents the importance of *multidimensional* human capital accumulation for individual career dynamics. Intuitively, experience in a given domain (e.g., industry, occupation, or task) disproportionately increases productivity in activities related to that domain. This richer view of human capital has proven useful for explaining observed patterns of job mobility and earnings dynamics (Gathmann and Schönberg (2010), Lise and Postel-Vinay (2020)). I contribute to this literature by providing evidence that workers accumulate domain-specific human capital within occupational career ladders.

Second, I relate to the empirical literature that has shown how automation and routine-biased technological change have driven large *economy-wide* shifts in the task content of work and the labor market returns to different types of skills over the past several decades (e.g., Acemoglu and Autor (2011), Atalay et al. (2020), Autor et al. (2003), Autor et al. (2006), Goos et al. (2014)). However, a common approach in this literature is to frame comparative advantage across tasks as either a fixed attribute of workers or one that can be mapped to a set of demographic characteristics, which often includes age. Empirically, I show that individuals specialize in different types of tasks at different stages of their careers, and these tasks need not be uniformly exposed to automation. My recovered learning estimates also suggest that comparative advantage across tasks evolves meaningfully with the specific labor market experiences that individuals accumulate early in their careers, rather than age alone.

Third, the model in this paper builds directly on existing work that develops dynamic models of occupational choice in order to study labor market adjustments in response to trade and technology shocks (Dix-Carneiro (2014), Dvorkin and Monge-Naranjo (2019), Humlum (2021), Traiberman (2019)). Two recent papers in this literature are closely related. First, Dvorkin and Monge-Naranjo (2019) develop a dynamic model of occupational choice with endogenous human capital accumulation decisions and task assignment between workers and machines. My model adopts their tractable microfoundation for within-occupation automation. However, their focus is primarily on the importance of mobility frictions across broad occupational categories for the effects of technological change on cross-sectional wage inequality. Consequently, their analysis

abstracts from ex ante heterogeneity across workers and life-cycle dynamics of human capital accumulation. In contrast, my model explicitly centers around the initial career choices of young workers and the structure of on-the-job learning within and across occupational career ladders. Quantitatively, I study how the effects of automation vary across workers with different skill sets and evolve over the transition.

Second, Althoff and Reichardt (2026) also develop a dynamic model of occupational choice with on-the-job skill accumulation. In their model, workers accumulate skills along a small number of dimensions and endogenously sort across occupations with different task and skill requirements, which together with a worker's current skills, determine their skill accumulation. I pursue an alternative approach that allows the productivity gains from occupational experience to vary flexibly across origin and destination occupation pairs. While this flexibility comes at the cost of lost parsimony, it allows me to characterize detailed occupational career ladders without specifying the underlying skills or mechanisms by which skills accumulate. Their counterfactual analysis also emphasizes two additional margins of AI-driven technological change, augmentation and simplification, that are absent from this paper.

Finally, this paper relates to a growing literature that studies the potential effects of AI-driven automation on early-career workers. Recent theoretical work has focused on how automation interacts with learning, highlighting potential inefficiencies arising from learning spillovers (Afrouzi et al., 2026; Ide, 2026), trade-offs between short-run productivity gains and long-run skill formation (Asriyan et al., 2026), and implications for organizational design (Friebel et al., 2026; Garicano and Rayo, 2026). My contribution to this literature is twofold. Empirically, I provide evidence of occupational career ladders along which workers move toward less AI-exposed occupations. Conceptually, I show that automation in early-career occupations can encourage further automation higher up the career ladder, by reducing the supply of workers skilled in those occupations. Related empirical studies analyze the employment effects of generative AI on junior and senior workers within occupations (Brynjolfsson et al., 2026; Hosseini Maasoum and Lichtinger, 2026b). While data limitations lead me to focus on occupation-level automation, the mechanisms present in this paper may also contribute to longer-run dynamics of automation of tasks within occupations.

Outline. The rest of this paper is organized as follows. Section 2 introduces the job profiles data and presents descriptive patterns of life-cycle occupational mobility that motivate key features of the quantitative model and the quantitative experiments. Section 3 introduces the model and Section 4 derives comparative statics that provide intuition for how on-the-job learning shapes the long-run impacts of labor-automating technologies within versus across career paths. Section 5 describes the procedure that I use to recover the parameters governing the labor supply model from data on occupational choices and cross-sectional wages. Section 6 then uses the model to quantitatively evaluate the potential labor market impacts of generative AI-driven automation shocks. Section 7 concludes.

2 Data and motivating evidence

In this section, I introduce the job profiles data and provide a detailed overview of my sample construction, given the relative novelty of this data. I use this dataset in two ways. First, I use it to document descriptive patterns of early career occupational mobility that show how the types of tasks that workers perform evolve as they gain labor market experience. These findings motivate the ingredients of the model that I develop in Section 3. Second, in Section 5, I use this data to empirically discipline key elements of the model.

2.1 Data source and sample construction

My primary data source is a résumé-based dataset with information on LinkedIn job profiles, provided by Revelio Labs through WRDS. This dataset has grown in popularity in economics research in recent years, as a rich source of information on firm- and worker-level employment patterns (Dorn et al. (2026), Hampole et al. (2025), Hosseini Maasoum and Lichtinger (2026b)). The main advantage of the Revelio data for my analysis, relative to other commonly used datasets with information on the occupations held by workers in the United States, such as the Current Population Survey (CPS) and the NLSY, is that it allows me to measure individual occupational transitions both over a long time horizon and at a granular level, given its large sample size. Precisely because college-educated workers and those employed in professional occupations are over-represented on LinkedIn, it is also a potentially useful source of data for learning about the career dynamics of the individuals that work in the occupations that are commonly considered to be most exposed to generative-AI driven automation.

Initial sample. The Revelio data contains information for each position and educational degree that a user lists on their (public) LinkedIn profile. A position is associated with a start date, an end date, an employer, a job title, and a job description. To construct an initial sample of eligible workers, I consider all users that (i) held an active position in the United States in 2022 and (ii) report a degree on their profile. I then restrict to all U.S.-based positions held by these users up to the end of 2022. I end the sample in 2022 in order to study preexisting features of career dynamics before the emergence of generative AI.

Measuring age and experience. I construct measures of age and experience based on the educational information reported on users' profiles, largely following the procedures described in Dorn et al. (2026). To construct potential labor market experience, I define the beginning of a worker's career as the end of their earliest dated undergraduate degree, including associate degrees. For workers that list overlapping or consecutive college degrees, I consider the latest graduation date of the combined educational spell, so that the career start date for a worker who transfers directly from an associate's degree program to a bachelor's degree program is the end of the bachelor's degree. For users that only list a high school degree on their profile, I define the start of their career as their high school graduation date. Similarly, I infer birth year from the user's earliest dated education episode. In particular, I assume that users are aged 18 at high school completion, 18 at bachelor's-program entry, and 19 at associate-degree-program entry.[2] Throughout, I restrict to positions held while workers are between the ages of 18 and 60. I also limit the sample to users with less than thirty years of potential experience in 2022 to limit the impact of older cohorts of LinkedIn users, for whom participation may be more selective and retrospective career histories less complete.

Worker panel. From the sample of eligible positions, I convert each user's work history into an annual panel, where an observation corresponds to the main job held by a user in a given year. Lastly, I define actual work experience as the number of years since (inferred) labor market entry that the user has held a main job. Additional details on the data cleaning and sample construction are described in Appendix A.2.

Figure A.1 reports the resulting number of users and observations, as I restrict the sample to users that entered the labor market after a minimum cutoff year (referred to as the "entry floor" year in the legend)

[2]Assumed ages are based on median gaps between high school end dates and college start dates. Missing education dates are imputed using median degree-specific program durations in the Revelio data. I require all degrees to have either a start or end date listed, but do not require both. However, all results are robust to excluding the small number of users with imputed career start dates.

and are observed holding a main job at least x years after their inferred labor market entry year (referred to as the "minimum experience cutoff" in the x-axis). In the descriptive regressions below, I restrict the sample to users that entered the labor market in 2010-2012, are observed in a main job at least ten years later, and that held a main job during the first five years of their career.[3] Results are qualitatively similar for cohorts that entered the labor market before 2010. Importantly, for these earlier cohorts, I also find that the within-individual changes in occupational characteristics that I document in this section slow down significantly after the first decade of work experience.

Assigning occupation codes. To map self-reported job titles to Standard Occupational Classification (SOC) codes, I use the occupation codes assigned by Revelio's proprietary classification algorithm, following existing work (Dorn et al. (2026), Hampole et al. (2025), Hosseini Maasoum and Lichtinger (2026b)).[4] However, I also refine some of these assignments based on several audits of the underlying job titles, employers, and education histories to correct mis-coded occupations and reduce spurious transitions, as described in Appendix A.2. I exclude employment in military occupations (SOC codes beginning with 55).

2.2 Motivating evidence

Life-cycle changes in task and skill content. I first document that young workers quickly transition to occupations with higher average salaries and non-routine task content in the first decade of their careers. Figure 1 plots the estimated coefficients on work experience from separate regressions of various occupation-level characteristics on three-year bins of work experience:

$$\text{Occupation characteristic}_{o(it)} = \alpha + \sum_{e} \beta_e \cdot \mathbb{1}[\text{Experience}_{it} = e] + \gamma_i + \varepsilon_{it}, \tag{1}$$

where γ_i denotes an individual fixed effect and $o(it)$ denotes the occupation o that user i holds in year t.[5] The set of occupation characteristics that I consider are the log of the average salary from the 2022 Occupational Employment and Wage Statistics (OEWS) and the five measures of non-routine and routine task content constructed in Acemoglu and Autor (2011).[6] Figure 1 shows that there are two important changes in the occupational sorting patterns of workers as they gain labor market experience. The first is a shift away from occupations intensive in manual tasks, towards occupations intensive in cognitive and interpersonal tasks. The second is a transition towards occupations with a higher non-routine task content. This latter finding is in line with recent work by Deming (2025), who shows that an alternative measure of occupational routineness falls quickly during the first several years of college-educated workers' careers, using data on the NLSY 79 and 97 cohorts. Notably, I find that these basic features of early career occupational mobility are present across users with different types of educational backgrounds.

In Figure A.2, I show how membership in broad occupational categories evolves with labor market experience, with the most pronounced life-cycle shifts representing transitions into managerial work. Finally, Figure A.3 reports the cumulative number of occupations and firms attached to the positions held by users during the

[3] Beyond these restrictions, I do not explicitly require continuous labor force attachment. However, measured work experience closely tracks potential experience in my sample of users.

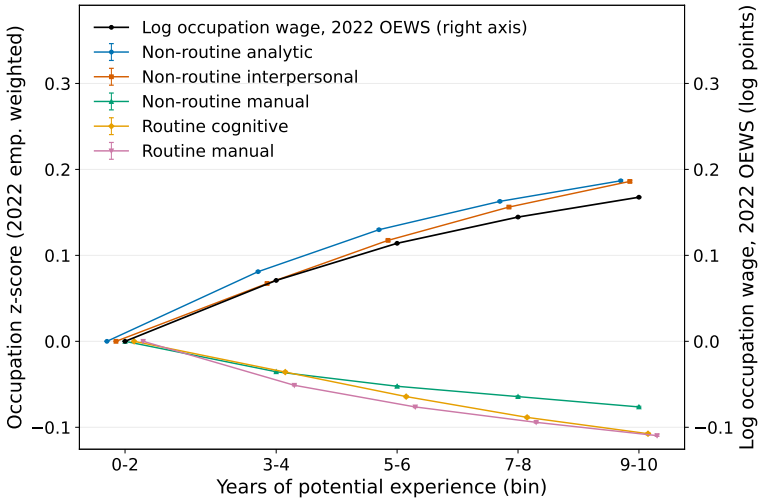
[4] Revelio assigns positions to standardized roles in its proprietary taxonomy and provides a unique mapping from their roles to 2019 O*NET-SOC occupations, which can be directly mapped one-to-one to detailed 2018 SOC codes.

[5] The results are essentially unchanged when Regression (1) includes potential experience, instead of work experience.

[6] The definitions of these occupational routineness measures are reported in Appendix Table A.1. I update the scores of Acemoglu and Autor (2011) using O*NET version 27.2.

first decade of their careers. The figure shows that the pace of occupational mobility slows considerably by the end of the first decade of labor market experience. Importantly, this slowing in occupational mobility persists even when I restrict the sample to users who later switched firms, suggesting that this finding is not simply an artifact of users ceasing to update their profiles.

Figure 1: Occupational skill and task content



Notes: Figure plots the estimated β_e coefficients from Regression (1). Standard errors are clustered at the user level, and 95% confidence intervals are shown.

Life-cycle changes in automation exposure. Next, I study how these patterns of early career mobility translate to changes in automation exposure to generative AI. To construct occupation-level measures of automation exposure, I average binary task-level automation exposure scores within an occupation, where tasks are the set of activities performed by workers in an occupation, as defined in O*NET.[7] I define a task as exposed if it is rated as fully or almost fully automatable in Eloundou et al. (2024), consistent with the recent approaches of Fan (2025) and Freund and Mann (2026). Appendix A.3.2 provides additional details on how Eloundou et al. (2024) assign task-level exposure scores and reports the distribution of occupational exposure scores and the relationship between automation exposure and average wages at the broad occupation level. In general, the most exposed categories of occupations are computer and mathematical, business and financial, and office and administrative support occupations.

The left panel of Figure 2 plots average exposure among users with 0-2 years of work experience, where I distinguish users by the level of educational attainment used to define their career start date.[8] The figure shows that early career exposure is highest among users with a bachelor's degree, consistent with their higher representation in professional occupations. More interestingly, the right panel of Figure 2 plots the estimated coefficients on experience from regressions of occupation-level automation exposure on work experience:

$$\text{Automation exposure}_{o(it)} = \beta + \sum_{e} \beta_e \cdot \mathbb{1}[\text{Experience}_{it} = e] + \alpha_i + \varepsilon_{it}, \tag{2}$$

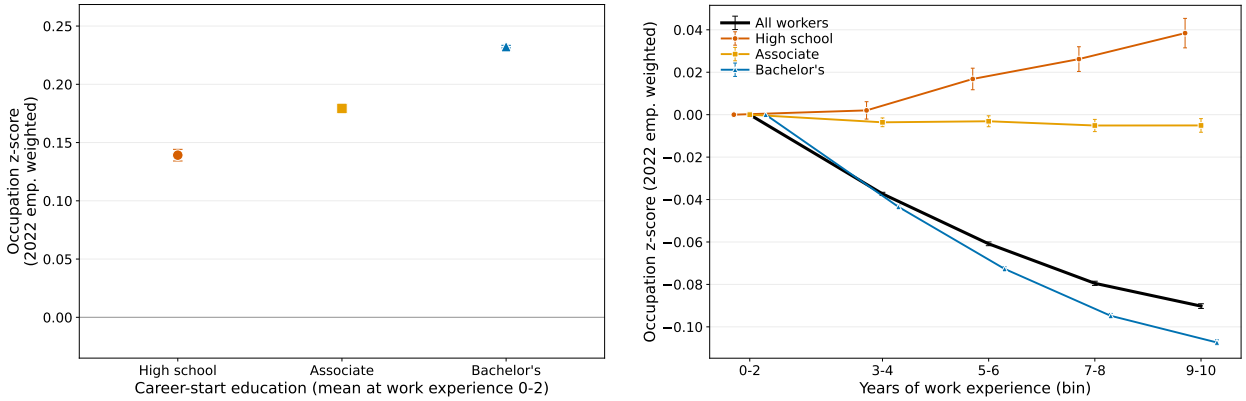
[7] I weight "core" tasks, as defined by O*NET, twice as much as non-core tasks, following Eloundou et al. (2024).

[8] Average exposure among users on LinkedIn is higher than in the general population, due to the over-representation of users in professional occupations. I find a similar ranking of initial exposure across young workers with different educational backgrounds in the ACS, as well as similar life-cycle patterns based on cross-sectional exposure profiles.

which I run separately for each group of users. The results show that, on average, automation exposure tends to fall with work experience among users that complete a bachelor's degree, while it rises or remains flat for those that complete only a high school or associate's degree. Given the inherent uncertainty in AI capabilities and limitations of occupation-level exposure scores, I view these findings as suggestive evidence that the occupational transitions that workers with a bachelor's degree tend to make early on in their careers reduce their average exposure to generative AI-driven automation.

The "Associate's" and "Bachelor's" samples also include users who later re-enrolled in higher degrees. Figure A.7 shows that the results are similar when restricting to users whose highest degree record matches the degree used to determine their labor market entry date.[9] Additionally, since the scores from Eloundou et al. (2024) are based on GPT-4 model capabilities, Figure A.8 reports results based on versions of these scores that have been updated for GPT-5 model capabilities, as constructed by Hosseini Maasoum and Lichtinger (2026a).[10] While the decline in average exposure for bachelor's degree holders is still present, it is much smaller in magnitude. In contrast, the increase in average automation exposure for users with less than a bachelor's degree becomes much more pronounced.

Figure 2: Automation exposure to generative AI



(a) Average exposure at 0-2 years of experience

(b) Changes in exposure

Notes: Panel (a) reports average exposure for the three educational groups at 0-2 years of work experience. Panel (b) plots the estimated β_e coefficients from Regression (2), where the regression is run separately for each educational group. Standard errors are clustered at the user level, and 95% confidence intervals are shown.

Examples of typical career paths. Finally, I exploit the size of the LinkedIn data to provide examples of the typical career trajectories that contribute to the decline in average automation exposure among bachelor's degree holders. To this end, I construct an occupational transition network at a 10-year horizon, focusing on the sample of workers that report a bachelor's degree on their profile and had 0-9 years of potential experience by 2010-2012. For each user in this sample, I then record their ultimate occupation, ten years later.[11] Figure 3 provides a visual representation of the resulting transition network. In Figure 4, I report the implied change in automation exposure for the top ten off-diagonal pairs in the resulting network, where

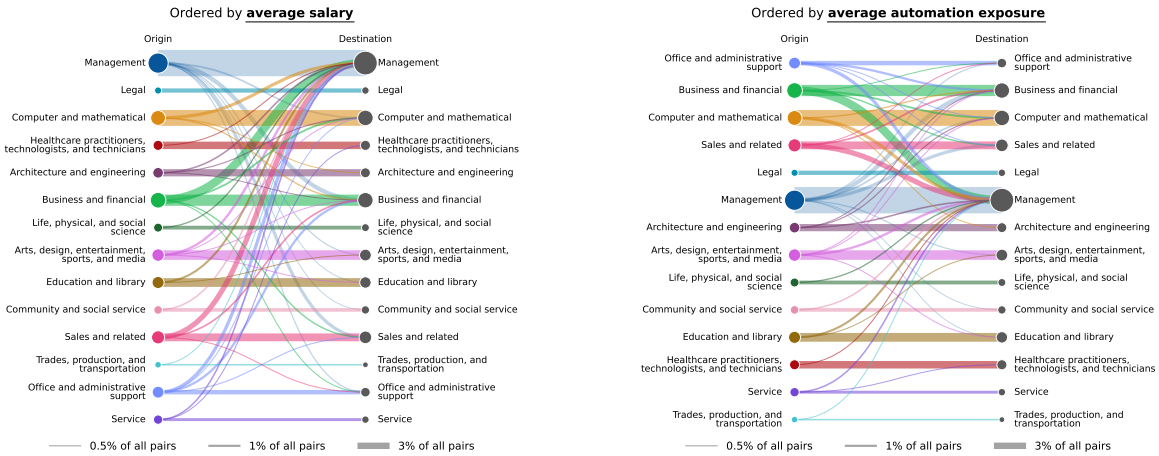
[9] Users whose labor market entry dates are determined by high school graduation are terminal by construction.

[10] There are some differences in the rubrics used by Eloundou et al. (2024) and Hosseini Maasoum and Lichtinger (2026a) to assign task-level exposure, so the scores are not a direct comparison of model capabilities.

[11] For this exercise, I also aggregate occupations to a broader classification, reported in Table A.2. I compute transitions at a ten-year horizon, rather than iterating an annual transition matrix, to capture longer-run occupational mobility, including intermediate switches, without assuming that annual transition probabilities are constant across career stages or depend only on current occupation.

I rank pairs based on the total mass of movers. Notably, the most common occupational transitions tend to include moves *within similar domains*. These moves include upward transitions (e.g., "financial analysts, advisors, and other financial specialists" to "financial managers" and "computer occupations" to "computer and information systems managers"), as well as horizontal transitions (e.g., "computer occupations" to and from "engineers and architects"). In addition, the figure shows that many of the most common transitions in the data also tend to reduce automation exposure to AI.

Figure 3: Visual representation of the 10-year transition network for users with a bachelor's degree



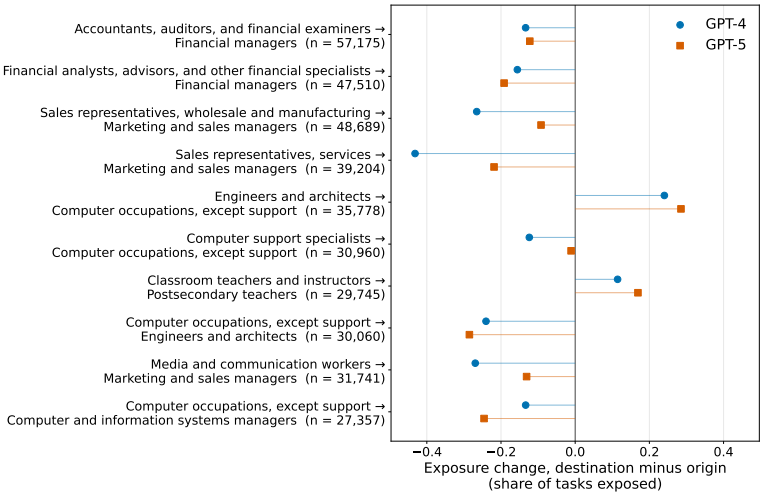
(a) Ranking occupations by average wages (2022 OEWS)

(b) Ranking occupations by average exposure

Notes: Both panels show the same observed ten-year flows between broad occupational groups among users with a bachelor's degree or higher. Origins are observed at zero to nine years of potential experience, and destinations ten years later. Links connect origin groups on the left to destination groups on the right; colors identify the origin group, and line widths are proportional to each flow's share of all observation pairs, including stayers. Only between-group flows accounting for at least 0.3% of all pairs are displayed. Node areas are proportional to each group's share of origins or destinations, including within-group stays. Groups are ordered from highest to lowest average annual wage in panel (a) and GPT-4-based automation exposure in panel (b). Both group averages use 2022 OEWS employment weights across detailed occupations.

Takeaways. Taken together, these patterns and examples offer suggestive evidence that (i) the task content of work changes meaningfully with labor market experience, (ii) workers differ both in their initial levels of exposure, as well as their life-cycle changes in exposure to generative AI-driven automation, (iii) occupations *within* common career paths are differentially exposed to AI-driven automation, and (iv) *domain-specific* experience is an important feature of professional career paths. In the next section, I develop a structural model in order to study the long-run impacts of heterogeneous automation shocks to occupations within and across career paths.

Figure 4: Changes in exposure for the most common occupational transitions in ten-year transition network



Notes: Figure shows the change in exposure for the top ten pairs of occupation transitions among LinkedIn users with a bachelor's degree and zero to nine years of potential experience in 2010-2012. GPT-4 refers to occupation-level automation exposure scores, computed using the task-level exposure scores of Eloundou et al. (2024), which are based on GPT-4 era model capabilities. GPT-5 refers to occupation-level automation exposure scores, computed using the task-level exposure scores of Hosseini Maasoum and Lichtinger (2026a), which are based on GPT-5 era model capabilities.

3 A model of on-the-job learning and automation

This section develops a dynamic general equilibrium model of the labor market with overlapping generations of workers that live and work for two periods. In the model, young workers enter the labor market with heterogeneous skills across occupations and make dynamic discrete choices about which occupation to work in. The evolution of their skill in each occupation depends on the occupation performed when young and is the only reason that the initial occupation choice is a dynamic decision problem. In Section 4 I derive comparative static results that provide intuition for how the structure of on-the-job learning shapes the long-run impacts of uneven occupation-level improvements in labor-automating technologies on wages and *observed automation* across occupations.

In these next two sections, I focus on a simplified model environment with ex ante homogeneous workers in order to highlight the novel mechanisms in the model and ease the notational burden. In the empirical implementation of the model outlined in Section 5, I explicitly incorporate ex ante heterogeneity across workers, motivated by the empirical evidence in the previous section which suggested that the impacts of AI-driven automation are likely to vary across different groups of workers. Similarly, while I focus on long-run outcomes in the theoretical analysis, in the quantitative experiments of Section 6 I also study the transitional dynamics of the economy following an automation shock, given the computational tractability of the model.

3.1 Model environment

Time is discrete and runs forever. Throughout, I use bold font to denote matrices and vectors, and I refer to "jobs" and "occupations" interchangeably. I focus on stationary equilibria without aggregate growth or uncertainty and therefore omit time subscripts and individual expectations of aggregate variables.

Demographics. The economy is populated by a constant mass of workers that live and work for two periods. At any given time, there are two age groups present in the labor market, referred to as young y and old o, each with unit mass. Upon first entering the labor market, workers are ex ante homogeneous.

Preferences. All workers have log preferences over consumption of the final good and receive an additive non-wage utility value from working in each occupation j, denoted by a_j. Young workers discount future payoffs with subjective discount factor $\beta \in (0,1)$. Flow utility in each period from consuming c units of the final good and working in occupation j is thus given by:

$$u(c,j) = \log(c) + a_j.$$

Jobs, skills, and learning. There are J types of jobs in the economy, indexed by $j \in \{1,...,J\}$. Workers enter the labor market with a vector of skills for each job type $\boldsymbol{\psi}^y = (\psi_1^y, \ldots, \psi_J^y) \in \mathbb{R}_{++}^J$. In each period, worker productivity in each occupation is subject to an idiosyncratic productivity shock, $\boldsymbol{\epsilon} \in \mathbb{R}_{++}^J$ that is drawn from a Fréchet distribution with unit scale and shape parameter $1/\nu$. These productivity shocks are assumed to be fully observed by all agents (firms and workers) at the beginning of each period. I additionally assume that they are independent across occupations, workers, and time and are fully transitory.[12] The effective labor that a worker with current skills $\boldsymbol{\psi}$ and productivity shock $\boldsymbol{\epsilon}$ can supply to occupation j is

$$\psi_j \epsilon_j.$$

Skill accumulation is passive, deterministic, and depends on the job performed when young. In particular, an old worker's skill in job k after first working in job j is given by:

$$\gamma_{jk} \cdot \psi_k^y, \quad \gamma_{jk} > 0.$$

In Section 3.3, I discuss how this reduced-form learning technology nests common formulations of on-the-job human capital accumulation in the literature.

Technology. A perfectly competitive, representative firm combines efficiency units of labor h_j with occupation-specific automation capital k_j in a nested CES aggregator to produce the final good Y (numeraire):

$$Y(\{h_j,k_j\}_j) = \left(\sum_j \alpha_j^{1/\lambda} y_j(h_j,k_j)^{\frac{\lambda-1}{\lambda}}\right)^{\frac{\lambda}{\lambda-1}}, \quad y_j(h_j,k_j) = \left((A_j^H h_j)^{\frac{\eta-1}{\eta}} + (A_j^K k_j)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}, \tag{3}$$

where $\sum_j \alpha_j = 1$ and A_j^H and A_j^K denote the productivities of labor and capital used in occupation j, respectively. $\lambda > 0$ is the elasticity of substitution between occupation-specific outputs $\{y_j\}$, and $\eta > 0$ is the elasticity of substitution between capital and labor in each occupation, which I assume to be greater than $\max\{1,\lambda\}$.[13] Capital is an intermediate good that can be transformed from the final good at the exogenous rate $1/q_j$, as in Acemoglu and Restrepo (2022). I abstract from the dynamics of capital accumulation in

[12] Fully transitory productivity shocks are a strong assumption that affords significant tractability. The empirical implementation in Section 5 accounts for persistent skill heterogeneity on the basis of observable characteristics.

[13] As discussed below, this assumption allows me to interpret improvements in $A_j^K q_j$ as labor-automating technological change.

order to focus on the dynamic role of labor supply.[14]

Microfoundation for automation. As shown by Dvorkin and Monge-Naranjo (2019) and proved in Appendix B.1, the occupation-level aggregator $y_j(h_j, k_j)$ in (3) can be microfounded as a task-level production function, in which there is a continuum of tasks involved in producing each occupation's output, capital and labor are perfect substitutes at the task level, and the task-level productivities of capital and labor governing comparative advantage across tasks are distributed according to a Fréchet distribution with scale parameters A_j^H and A_j^K and common shape parameter $\eta - 1$, consistent with the restriction that $\eta > 1$.[15]

Labor and asset markets. The labor market for efficiency units of labor in each occupation is perfectly competitive. A worker with current skills $\boldsymbol{\psi}$ and productivity shock $\boldsymbol{\epsilon}$ earns $w_j \psi_j \epsilon_j$, where w_j is the price of an efficiency unit of labor in occupation j, determined in equilibrium. There are no financial markets for workers to borrow or save, and all workers consume hand-to-mouth.[16]

3.2 Optimality conditions and equilibrium

Old worker occupational choices. Given occupational skill prices $\boldsymbol{w}$, an old worker with current skills $\boldsymbol{\psi}$ and productivity shock $\boldsymbol{\epsilon}$ chooses their occupation to maximize flow utility:

$$e^o(\boldsymbol{w}, \boldsymbol{\psi}, \boldsymbol{\epsilon}) \coloneqq \arg\max_j \log(w_j \psi_j \epsilon_j) + a_j.$$

The share of old workers with skills $\boldsymbol{\psi}$ choosing to work in occupation j and the expected utility of old workers with skills $\boldsymbol{\psi}$ are then given by their familiar logit forms[17]:

$$e_j^o(\boldsymbol{w}, \boldsymbol{\psi}) = \frac{(w_j \psi_j)^{1/\nu} \exp(a_j/\nu)}{\sum_k (w_k \psi_k)^{1/\nu} \exp(a_k/\nu)}, \tag{4}$$

$$\bar{V}^o(\boldsymbol{w}, \boldsymbol{\psi}) = \nu \log \left[\sum_k (w_k \psi_k)^{1/\nu} \exp(a_k/\nu) \right] + \nu \gamma_E. \tag{5}$$

Let $\boldsymbol{\gamma}_{\ell\cdot}$ denote row ℓ of $\boldsymbol{\gamma}$ and $\odot$ denote the element-wise vector product. Notationally, it is helpful to define $e_j^{o|\ell}(\boldsymbol{w}) \coloneqq e_j^o(\boldsymbol{w}, \boldsymbol{\gamma}_{\ell\cdot} \odot \boldsymbol{\psi}^y)$ as j's employment share among old workers who previously worked in ℓ and $\bar{V}^{o|\ell}(\boldsymbol{w}) \coloneqq \bar{V}^o(\boldsymbol{w}, \boldsymbol{\gamma}_{\ell\cdot} \odot \boldsymbol{\psi}^y)$ as the expected value of old workers that initially worked in ℓ.

Young worker occupational choices. Young workers choose their initial occupation to maximize *expected lifetime utility*:

$$e^y(\boldsymbol{w}, \boldsymbol{\epsilon}) = \arg\max_j \log(w_j \psi_j^y \epsilon_j) + a_j + \beta \bar{V}^{o|j},$$

where I have suppressed the dependence of the employment choices of young workers on $\boldsymbol{\psi}^y$, which is common to all workers. Young worker employment choices $e_j^y(\boldsymbol{w})$ and ex ante values $\bar{V}^y(\boldsymbol{w})$ thus take on a similar

[14] Given that I am focused on long-run steady states, the main object of interest in a model with endogenous capital accumulation would be the steady state cost of capital.

[15] In principle, the dispersion of these task-level productivities could vary across occupations, yielding heterogeneous elasticities of substitution between capital and labor across occupations. However, since I do not have sufficient data to estimate occupation-specific elasticities between capital and labor, in the empirical application I calibrate η to a single value and explore sensitivity of the results to this choice.

[16] I discuss the implications of this restriction in Section 3.3.

[17] γ_E denotes the Euler–Mascheroni constant.

form to the expressions for old workers, except they include the continuation values associated with working in each occupation when young, which reflect differences in learning opportunities across occupations:

$$e_j^y(\boldsymbol{w}) = \frac{\exp(\log(w_j\psi_j^y) + a_j + \beta\bar{V}^{o|j})^{1/\nu}}{\sum_k \exp(\log(w_k\psi_k^y) + a_k + \beta\bar{V}^{o|k})^{1/\nu}} \tag{6}$$

$$\bar{V}^y(\boldsymbol{w}) = \nu \log\left[\sum_k \exp(\log(w_k\psi_k^y) + a_k + \beta\bar{V}^{o|k})^{1/\nu}\right] + \nu\gamma_E. \tag{7}$$

To gain intuition for the roles of β and ν, note that when $\beta = 0$, young and old workers would make the same employment choices, *conditional on skill*, as young workers would no longer value learning opportunities. With on-the-job skill accumulation, however, old and young workers would still generally make different employment choices due to differences in their accumulated skills. Additionally, ν governs the dispersion of the idiosyncratic productivity shocks and the strength of sorting on the common component of occupational payoffs. When $\nu = 0$, all observed sequences of occupation choices must offer the same present discounted value of utility.[18] With $\nu > 0$, differences in realized productivity shocks translate to differences in utility, and young workers make heterogeneous occupational choices due to differences in their initial productivity draws. These initial occupational choices in turn shape their skill accumulation. As a result, old workers make heterogeneous employment decisions due to differences in their accumulated skills and their idiosyncratic productivity shocks in the second period.

Aggregation. Total employment in occupation j, which I denote by $e_j(\boldsymbol{w})$, is given by the sum of the employment of young and old workers in j:

$$e_j(\boldsymbol{w}) = e_j^y(\boldsymbol{w}) + \sum_\ell \underbrace{e_\ell^y(\boldsymbol{w}) \cdot e_j^{o|\ell}(\boldsymbol{w})}_{\ell \to j \text{ workers}}. \tag{8}$$

The second term is the employment contribution of old workers to occupation j, and it sums over all initial occupations ℓ, reflecting that the employment decisions of young workers shape the employment decisions of old workers, via their accumulated skills. Similarly, total effective labor supply in each occupation j, denoted by $h_j^S(\boldsymbol{w})$ is given by the sum of the effective labor supply of young and old workers:

$$h_j^S(\boldsymbol{w}) = \underbrace{\psi_j^y \cdot \mathbb{E}[\epsilon_j | j = e^y(\boldsymbol{w}, \boldsymbol{\epsilon})] \cdot e_j^y(\boldsymbol{w})}_{:= h_j^y} + \sum_\ell \underbrace{\psi_j^y \gamma_{\ell j} \cdot \mathbb{E}[\epsilon_j | j = e^{o|\ell}(\boldsymbol{w}, \boldsymbol{\epsilon})] \cdot e_\ell^y(\boldsymbol{w}) \cdot e_j^{o|\ell}(\boldsymbol{w})}_{:= h_j^{o|\ell}}, \tag{9}$$

where $\gamma_{\ell j}$ in the second term reflects the direct effect of skill accumulation on the effective labor supply of old workers. Given the distributional assumptions on $\boldsymbol{\epsilon}$, h_j^y and $h_j^{o|\ell}$ can be written as:

$$h_j^y(\boldsymbol{w}) = \Gamma(1-\nu) \cdot \psi_j^y \cdot (e_j^y(\boldsymbol{w}))^{1-\nu} \tag{10}$$

$$h_j^{o|\ell}(\boldsymbol{w}) = \Gamma(1-\nu) \cdot \psi_j^y \gamma_{\ell j} \cdot e_\ell^y(\boldsymbol{w}) \cdot (e_j^{o|\ell}(\boldsymbol{w}))^{1-\nu}, \tag{11}$$

where $\Gamma(\cdot)$ denotes the Gamma function.[19] The $1-\nu$ exponent on the employment shares of young and old workers captures how selection on idiosyncratic productivity shapes effective labor supply: when only a small number of workers work in j, those workers must have received unusually large productivity draws.

[18] Section 4.2 constructs an illustrative example that takes $\nu \to 0$.

[19] Throughout, I assume that $\nu < 1$ to ensure that the expected value of the idiosyncratic productivity shocks is finite.

Effective labor supply thus increases less than one-for-one with employment, as expansion draws in workers with less favorable productivity shocks.

Labor demand. The problem of the representative firm is to maximize output net of labor and capital costs, taking as given occupational skill prices:

$$\max_{\{h_j,k_j\}} Y - \sum_j w_j h_j - \sum_j k_j/q_j,$$

subject to the production function defined in (3). Cost minimization at the occupation-level implies a familiar expression for the share of income paid to labor in occupation j:

$$s_j^H \equiv \frac{w_j h_j}{w_j h_j + k_j/q_j} = \frac{(A_j^H/w_j)^{\eta-1}}{(A_j^H/w_j)^{\eta-1} + (A_j^K q_j)^{\eta-1}}. \tag{12}$$

Under the task-based microfoundation for $y_j(h_j, k_j)$, the labor share and the share of tasks assigned to labor in each occupation are identical, so that improvements in the efficiency of capital also lead to a higher share of tasks being assigned to capital. It is in this sense that I refer to improvements in $q_j A_j^K$ as *labor-automating technological change in* j and the capital share, $s_j^K = 1 - s_j^H$, as the *extent of automation* in j. The firm's full cost minimization problem yields the following inverse labor demand curve in occupation j:

$$w_j = \alpha_j^{1/\lambda} \cdot \left(s_j^H(w_j)\right)^{\xi} \cdot \left(\frac{Y}{h_j^D(\boldsymbol{w})}\right)^{1/\lambda} \cdot \left(A_j^H\right)^{\frac{\lambda-1}{\lambda}}, \quad \xi = \frac{\eta-\lambda}{(\eta-1)\lambda} > 0, \tag{13}$$

as shown in Appendix B.2. Analogous to Acemoglu and Restrepo (2022), this expression already suggests that improvements in labor-automating technology will have two effects on labor demand, holding fixed wages. First, there will be a positive productivity effect, arising from improvements in total output Y. Second, there will be a negative effect from the endogenous reassignment of tasks from labor to capital, showing up as a reduction in s_j^H. Section 4 formalizes this simple intuition. I next define a stationary competitive equilibrium of this economy.

Definition 1 (Stationary Equilibrium). *A stationary equilibrium consists of a collection of occupational skill prices* $\{w_j\}$*, factor demands* $\{k_j, h_j\}$*, and employment choices* $\{e^y(\boldsymbol{\varepsilon}), e^{o|\ell}(\boldsymbol{\varepsilon})\}$*, such that (i) given wages, employment choices* $\{e^y(\boldsymbol{\varepsilon}), e^{o|\ell}(\boldsymbol{\varepsilon})\}$ *are consistent with worker optimization, (ii) marginal products of capital and labor in each occupation are equated to their factor prices, (iii) the labor market for efficiency units of labor in each occupation clears* ($h_j^S = h_j^D$)*, and (iv) the ideal price condition is satisfied:*

$$1 = \sum_j \alpha_j p_j^{1-\lambda}, \text{ where } p_j^{1-\eta} \equiv (A_j^H/w_j)^{\eta-1} + (A_j^K q_j)^{\eta-1}.$$

3.3 Discussion

Learning technology. An advantage of the flexibility of $\boldsymbol{\gamma}$ is that it can capture several common formulations of on-the-job human capital accumulation in the literature. For example, if worker skills exhibit a common experience profile regardless of the initial or destination occupation, then $\gamma_{jk} \equiv g$. This restriction on $\boldsymbol{\gamma}$ resembles the standard life-cycle model of general human capital accumulation that accompanies labor

market experience. Alternatively, if skills for certain occupations grow at different rates, but do not depend on the occupation performed when young, then the rows of $\boldsymbol{\gamma}$ would be identical. This case captures the possibility that manual or managerial skills simply grow (or depreciate) at different rates over the life-cycle, independent of occupational experience.

In contrast, if the rate at which workers develop skills depends on the job performed when young, then the rows of $\boldsymbol{\gamma}$ will generally differ. For example, if workers accumulate only occupation-specific human capital on-the-job, then $\gamma_{jj} > 1$ and $\gamma_{jk} \leq 1$ for $j \neq k$.[20] If workers develop skills on-the-job that are valuable for other occupations, then $\gamma_{jk} > 1$ for some $k \neq j$. Importantly, this last case can capture occupational career ladders as well as transferability of skills across occupations. While passive on-the-job learning departs from the costly investment problem formulated in Ben-Porath (1967), young workers still make meaningful skill investment decisions in my model, via their initial job choices. In Section 5, I outline and implement a model-based approach to recover $\boldsymbol{\gamma}$ from data on occupational choices.

Automation. Starting directly from a CES production function in capital and labor departs from a more general task-based approach, as formulated in Acemoglu and Restrepo (2022).[21] However, an advantage of its simplicity is that it allows for a transparent analysis of the "ripple effects" of exogenous changes in technology *on automation*. As shown in the next section, exogenous improvements in capital efficiency in occupation j generally induce changes in occupational skill prices in general equilibrium. These changes in skill prices, in turn, shape the extent of automation in each occupation, which is summarized in the capital shares. Without putting structure on the task-level productivities of capital and labor, it is not possible to characterize the ripple effects of exogenous technology shocks on the extent of automation in occupations that are not directly subject to an automation shock.[22]

There are two limitations of the CES approach. First, it does not directly capture changes in the task-level *feasibility* of automation. In particular, $A_j^K \cdot q_j$ affects the productivity of capital in *all* tasks performed within an occupation, including those that are already assigned to capital. Hence, labor-automating technological change in this framework does not allow for a distinction between the intensive margin, capital-deepening effects of improvements in capital efficiency and the extensive margin, "displacement" effects of automation that are emphasized in Acemoglu and Restrepo (2019) and Acemoglu and Restrepo (2022). Second, it requires an estimate of the within-occupation, long-run elasticity of substitution between labor and automation capital, which may differ from standard measures of the elasticity of substitution between capital and labor.

Incomplete markets. The combination of log utility and the absence of financial markets for workers to smooth their consumption across periods and idiosyncratic states, while standard in dynamic discrete choice models of occupational choice, introduces an inefficiency into the model.[23] Importantly, this inefficiency is not essential for generating dynamic linkages in labor supply that transmit automation shocks across

[20] The learning technology thus also allows for depreciation of skills while working on unrelated tasks.

[21] Acemoglu and Restrepo (2022) show that their cost-minimized task-based production function has a reduced-form representation as a CES production function in labor and capital. The elasticity of substitution in the reduced-form representation (holding fixed task assignments or allowing them to adjust) need not be larger than one. Their estimation accounts for the effects of changes in equilibrium wages on industry-level capital shares.

[22] A common approach in recent studies is to assume that all automatable tasks are automated, shutting down the possibility of such ripple effects (Althoff and Reichardt (2026), Fan (2025), Freund and Mann (2026)). While this is a realistic assumption for studying short-run wage responses, it may miss some of the longer-run consequences of equilibrium labor reallocation.

[23] In ongoing work, I am studying the implications of this inefficiency for the occupational allocation and extent of automation in the initial steady state, as well as the impacts of changes in technology on this economy.

occupations. In particular, in Section 4.2 I consider a stylized example economy that generates the same qualitative wage and capital-share spillovers under linear preferences.

4 Equilibrium impacts of labor-automating technologies

I now characterize how wages and the extent of automation in each occupation respond to small changes in labor-automating technology in long-run equilibrium. It is helpful to first define the direct effect of small improvements in labor-automating technology in j on labor demand in occupation j, holding fixed wages and total output:

$$z_j \coloneqq d\log h_j^D\big|_{\boldsymbol{w},Y} = \lambda\xi \cdot d\log s_j^H\big|_{\boldsymbol{w}} = -(\eta-\lambda)s_j^K\left(d\log A_j^K + d\log q_j\right). \tag{14}$$

I refer to this change in labor demand as the *direct effect* of labor-automating technologies on labor demand in occupation j and denote it by z_j below. It is negative whenever $\eta > \lambda$, reflecting standard scale versus substitution logic: when occupations are less substitutable in production than capital and labor within occupation j, the direct effect of improvements in capital-augmenting productivity in j is to lower labor demand in j. Similarly, define the conditional elasticity of labor demand for j, denoted ε_j^d, as the magnitude of the change in labor demand in occupation j in response to small changes in w_j, holding technology and total output fixed:

$$\varepsilon_j^d \coloneqq -\frac{\partial\log h_j^D}{\partial\log w_j} = \eta\cdot s_j^K + \lambda\cdot(1-s_j^K) > 0. \tag{15}$$

Intuitively, the reduction in the demand for labor in occupation j in response to an increase in price of labor in j is increasing in both the substitutability of (i) factors within j and (ii) occupations. Finally, define the elasticity of labor supply in occupation j with respect to changes in wages in occupation k, $\varepsilon_{jk}^S \equiv \partial\log h_j^S/\partial\log w_k$. These objects are sufficient to characterize the equilibrium response of wages and capital shares in occupation j to an exogenous shock to labor-automating technology $\boldsymbol{A}^K\cdot\boldsymbol{q}$:

$$d\log w_j = \overbrace{\frac{d\log Y}{\varepsilon_j^d+\varepsilon_{jj}^S}}^{\text{Aggregate Productivity}} + \overbrace{\frac{z_j}{\varepsilon_j^d+\varepsilon_{jj}^S}}^{\text{Direct}} - \overbrace{\frac{\sum_{k\neq j}\varepsilon_{jk}^S\cdot d\log w_k}{\varepsilon_j^d+\varepsilon_{jj}^S}}^{\text{Cross-occupation spillovers in labor supply}} \tag{16}$$

$$d\log s_j^K = (\eta-1)\cdot s_j^H\cdot\left(\underbrace{d\log w_j}_{\text{Equilibrium Feedback}} + \underbrace{d\log A_j^K + d\log q_j}_{\text{Direct Automation}}\right). \tag{17}$$

Equation (16) shows that there are generically three effects of improvements in automation technologies on wages in equilibrium. The first is a standard aggregate productivity effect, which tends to raise wages in all occupations.[24] The second is the direct effect of improvements of labor-automating technology in j on labor demand in j. It is negative for occupations directly experiencing an automation shock, and zero otherwise. The third term captures the *ripple effects* that arise from the equilibrium reallocation of labor supply across occupations in response to changes in wages. All three terms are scaled by $\varepsilon_j^d+\varepsilon_{jj}^S$, reflecting how substitution across factors and occupations dampens (or potentially amplifies) uneven shocks to technology.

[24] Appendix B.3 provides derivations and additional technical details. It also derives the response of aggregate output and the aggregate labor share in terms of initial shares and elasticities, as well as analogous expressions for the equilibrium impacts of labor augmenting technologies.

Importantly, the first and last effects are present for all occupations, including those not directly experiencing an automation shock. Equation (17) clarifies that these equilibrium movements in wages will feed back into *observed automation*, even for the occupations that do not experience an automation shock.

Cross-occupation labor-supply spillovers describe how changes in skill prices in other occupations shift labor supply in occupation j. As usual, a wage increase in an occupation that is a labor-supply substitute for j draws labor away from j, placing upward pressure on its skill price. However, in the next section, I show that on-the-job learning can also generate *complementarities* in labor supply, which I define as $\varepsilon^S_{jk} > 0$ for $j \neq k$. As a consequence, an automation shock that lowers the relative skill price of a given occupation, instead lowers labor supply in its complementary occupations, putting upward pressure on their skill prices.

This channel operates alongside the common productivity effect of technological improvements.[25] The focus of my analysis is therefore primarily on how career linkages shape occupation-specific deviations in skill prices and automation *away* from this common response. After characterizing the behavior of labor supply in the general occupational choice model, in Section 4.2 I consider a stylized example economy with three occupations. This simple example illustrates how on-the-job learning can lead to complementarities in labor supply. It also shows how the effects of an automation shock on skill prices and capital shares across occupations depend on where the shock occurs within the career ladder.

4.1 Labor supply responses in the dynamic occupational choice model

The elasticity of labor supply in occupation j with respect to a change in the wage in occupation k, ε^S_{jk}, is given by the sum of three terms:

$$\varepsilon^S_{jk} = \underbrace{(1-\nu)\cdot\frac{h^y_j}{h_j}\cdot\frac{\partial \log e^y_j}{\partial \log w_k}}_{\text{Young reallocation}} + \underbrace{(1-\nu)\cdot\sum_\ell \frac{h^{o|\ell}_j}{h_j}\cdot\frac{\partial \log e^{o|\ell}_j}{\partial \log w_k}}_{\text{Old reallocation}} + \underbrace{\sum_\ell \frac{h^{o|\ell}_j}{h_j}\cdot\frac{\partial \log e^y_\ell}{\partial \log w_k}}_{\text{Dynamic composition effect}} . \tag{18}$$

The first term captures how the effective labor supply of young workers in occupation j responds to changes in wages in k. It includes the employment response of young workers and the implied change in selection (captured in the $1-\nu$ term), weighted by the contribution of young workers to labor supply in j, reflected in h^y_j/h_j. Similarly, the second term summarizes how the effective labor supply of old workers in occupation j responds, holding fixed the skill distribution of old workers at the initial equilibrium. Lastly, the third term captures a *dynamic composition effect*, which reflects how the reallocation of young workers reshapes effective labor supply in occupation j by changing the skill distribution of old workers. This term holds fixed the occupational choices of old workers at the initial equilibrium, reflected in the $h^{o|\ell}_j/h_j$ weights in the sum.

Focusing first on the old reallocation term, the logit choice model delivers standard substitution patterns for the employment responses of old workers:

$$\frac{\partial \log e^{o|\ell}_j}{\partial \log w_k} = \frac{1}{\nu}(\mathbb{1}[j=k] - e^{o|\ell}_k). \tag{19}$$

The responsiveness of old worker employment in j to changes in wages in k is thus governed solely by ν and

[25] In particular, Appendix B.3 shows that technological improvements generate a positive average log change in occupational skill prices, weighted by initial wage-bill shares. This insight reflects the assumption of a perfectly elastic capital supply, as discussed in Caselli and Manning (2019).

initial employment of old workers in occupation k, neither of which depend directly on j. The first source of complementarities in labor supply is the young reallocation term. To see this, note that the response of young worker employment in j to changes in wages in k is given by:

$$\frac{\partial \log e_j^y}{\partial \log w_k} = \frac{1}{\nu}(\mathbb{1}[j=k] - e_k^y) + \frac{\beta}{\nu} \cdot \left[\underbrace{e_k^{o|j}}_{d\bar{V}^{o|j}/d\log w_k} - \sum_\ell e_\ell^y \cdot \underbrace{e_k^{o|\ell}}_{d\bar{V}^{o|\ell}/d\log w_k} \right]. \tag{20}$$

The term in brackets reflects the fact that changes in w_k induce changes in the continuation value of working in each occupation ℓ when young, which depend directly on the likelihood of working in k when old.[26] Equation (20) implies that young worker employment in j increases with wages in k (for $k \neq j$) whenever the probability of directly working in k as a young worker is sufficiently low relative to the discounted increase in probability of working in k that working in occupation j induces.[27] Recall that on-the-job learning is the only way in which working in j shifts the likelihood of working in k when old. The second term is thus positive only when working in j tilts workers' future skill portfolio toward occupation k.

Equation (20) clarifies that the model implies a straightforward condition to check in the data for the presence of complementarities in young worker labor supply, conditional on a value of β. Whether a positive employment response of young workers is sufficient to induce a positive response at the occupation level, however, depends on the contribution of young workers to overall labor supply in j. Importantly, this complementarity disappears when young workers are not forward-looking or when learning is (log additively) separable in the origin and destination occupation.[28]

The second source of potential complementarities is the dynamic composition effect, which is the last term in the decomposition in Equation (18). Appendix B.4 shows that it can be re-written as:

$$\sum_\ell \frac{h_j^{o|\ell}}{h_j} \cdot \frac{\partial \log e_\ell^y}{\partial \log w_k} = \frac{\Gamma(1-\nu)\cdot \psi_j^y}{h_j} \cdot \mathrm{Cov}_{e_\ell^y}\left[\gamma_{\ell j} \cdot (e_j^{o|\ell})^{1-\nu}, \frac{\partial \log e_\ell^y}{\partial \log w_k} \right], \tag{21}$$

where Cov_{e^y} denotes the e_ℓ^y-weighted covariance operator. Equation (21) clarifies that the dynamic composition effect is positive whenever an increase in w_k induces an increase in young employment in the occupations ℓ that tend to lead to higher effective labor supply of old workers in occupation j. A positive composition effect can thus reflect occupational career ladders, as well as the transferability of skills across occupations with overlapping skill requirements. It disappears when on-the-job learning does not depend on the initial occupation (e.g., if skills vary only with age). However, unlike the complementarities in young worker labor supply, the dynamic composition effect survives when $\gamma_{\ell j} = g_\ell \cdot g_j$. This is because it includes the direct effect of young workers' occupational choices on the *level of human capital* of old workers. Reallocation of young workers toward occupations that provide better learning opportunities on average (higher g_ℓ) tends to raise future effective labor supply in j even if it does not tilt future employment choices toward j.

Conditional on values for ν and β, the three terms in Equation (18) can be computed from initial employment shares (e_j^y), occupational transition probabilities ($e_j^{o|\ell}$), and average wages by current occupation and age

[26] With $\beta < 1$, Appendix B.4 shows that the own-wage employment response of young workers always remains positive.

[27] This condition generalizes naturally to the case of multiple time periods. In particular, the relevant object is the discounted increase in future probability of working in occupation k, summed over all future time periods.

[28] To see the latter point, suppose $\gamma_{j\ell} = g_j \cdot g_\ell$. Equation (4) shows that working in occupation j does not affect the likelihood of choosing occupation ℓ when old, and the second term in Equation (20) thus disappears.

(and initial occupation for old workers). In particular, e_j^y and $e_j^{o|\ell}$ determine the employment elasticities in (19) and (20). Multiplying each group's average wage by its employment mass then gives its contribution to occupation j's total wage bill. Because all workers in j face the same skill price w_j, these wage-bill shares equal the effective-labor weights h_j^y/h_j and $h_j^{o|\ell}/h_j$. Measuring these objects in the data thus suffices to evaluate all three components of the occupation-level labor supply elasticities.

Takeaways. Job-dependent learning generates two potential sources of positive cross-wage elasticities in occupation-level labor supply. In Section 5.5, I directly evaluate these elasticities ε_{jk}^S in the quantified labor supply model. The exercise shows that complementarities in labor supply are not just a theoretical possibility, but an empirically relevant phenomenon. Applying the decomposition in Equation (18) to the model-implied elasticities of occupational labor supply, I find that the dynamic composition effect accounts for the largest cross-wage labor supply responses in the quantitative model, both positive and negative. This result reflects how accounting for endogenous skill accumulation decisions increases the long-run responsiveness of occupational labor supply to changes in relative skill prices.

4.2 Illustrative example with three occupations

In this section, I consider a simple economy with three occupations. Young workers enter the labor market with the initial skill vector $\boldsymbol{\psi}^y = (1, 1, s)$, where $s < 1$. In words, young workers are equally skilled in jobs one and two but initially less skilled in job three. The on-the-job learning technology is characterized by $\gamma_{2,3} = 1/s > 1$ and $\gamma_{j,k} = 1$ otherwise. The interpretation of this stylized learning technology is that job two is a "stepping-stone" or "training" occupation for job three: after working in job two, skills in job three catch up to skills in jobs one and two, which are assumed to be constant over time.

Figure 5 illustrates the behavior of labor supply under this simple learning structure. Panel (a) shows that young worker employment in job two *increases* in wages for job three, reflecting how forward-looking behavior, together with on-the-job learning, can lead to complementarities in young worker labor supply. Panel (b) confirms that old worker employment exhibits standard substitution patterns: conditional on first working in job two, old workers are less likely to work in job three as w_2 increases. Finally, panel (c) illustrates the dynamic composition effect: an increase in w_2 attracts more young workers into job two, raising labor supply in job three by increasing the share of workers skilled in job three. However, when w_2 is sufficiently high, static substitution forces dominates and old workers switch to job two, generating the non-monotonicity in job three's labor supply.

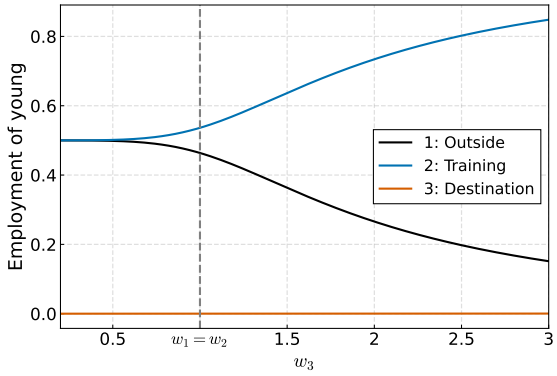
To analytically characterize how this economy responds to an automation shock, I consider the limit in which $\nu \to 0$.[29] I focus on stationary equilibria in which a positive mass of workers follows one of two career paths: working in occupation one (the "outside" occupation) at both ages, or working in occupation two when young and occupation three when old (the "training" career path). Let $e_2^y \in (0, 1)$ denote the initial share of young workers that enter the training occupation, so that $e_1^y = 1 - e_2^y$ and $e_3^y = 0$. In order for these career choices to be optimal, skill prices must satisfy:

$$w_3 > w_1 > w_2 > 0, \qquad sw_3 < w_1, \qquad \log w_2 + \beta \log w_3 = (1 + \beta) \log w_1.$$

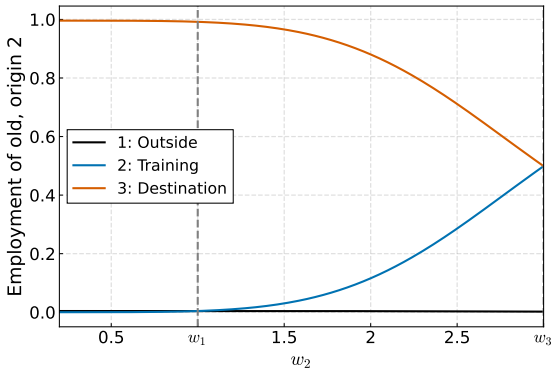
For simplicity, I assume that capital shares are *initially* common across occupations, $s_j^K = s^K \in (0, 1)$,

[29] Taking $\nu \to 0$ eliminates idiosyncratic productivity dispersion, simplifying the comparative statics.

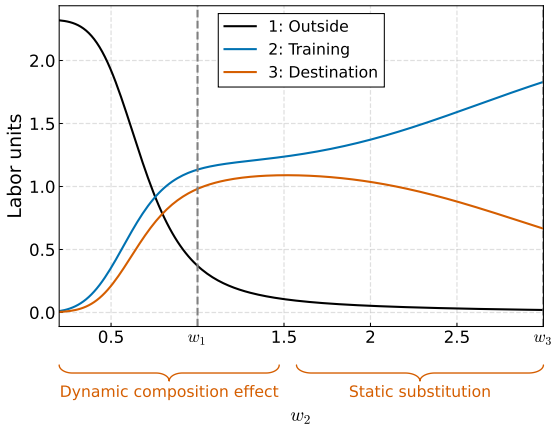
Figure 5: Decomposing dynamic complementarities in occupation-level labor supply



(a) e_j^y as w_3 varies



(b) $e_3^{o|2}$ as w_2 varies



(c) h_j as w_2 varies

Notes: Figure plots occupational employment and effective labor supply under the learning technology of Section 4.2, with log utility, $\nu = 0.2$, $\beta = 0.95^{20}$, and $s = 0.25$. Panel (a) varies the destination skill price w_3, holding $w_1 = w_2 = 1$. Panel (b) varies the training skill price w_2, holding $w_1 = 1$ and $w_3 = 2$, and shows old workers' occupational choice probabilities conditional on having worked in occupation two when young. Panel (c) varies w_2, holding $w_1 = w_3 = 1$, and shows effective labor supply. In panel (c), old workers' origin shares equal the young employment shares implied by each wage vector, incorporating the dynamic composition effect.

which implies a common conditional demand elasticity across occupations $\varepsilon_j^d = \varepsilon^d$. Appendix B.5 provides restrictions on the technology primitives that support this initial equilibrium and establishes that sufficiently small technology changes admit a unique nearby equilibrium in which both career paths remain active.

The next remark characterizes the first-order response of this economy to an exogenous automation shock in a single occupation, starting from an initial stationary equilibrium in which both career paths are active.

Remark 1 (Automation within vs across career paths). *Consider an automation shock* $\tau = d\log A_i^K + d\log q_i > 0$ *to occupation* i. *Let* $\Delta_i = d\log(w_3/w_1)$ *denote the resulting change in skill prices of job three relative to job one, given by:*

$$\Delta_1 = 0, \qquad \Delta_2 = -\Delta_3 = \frac{(\eta - \lambda)s^K}{\varepsilon^d(1+\beta)} \cdot \tau.$$

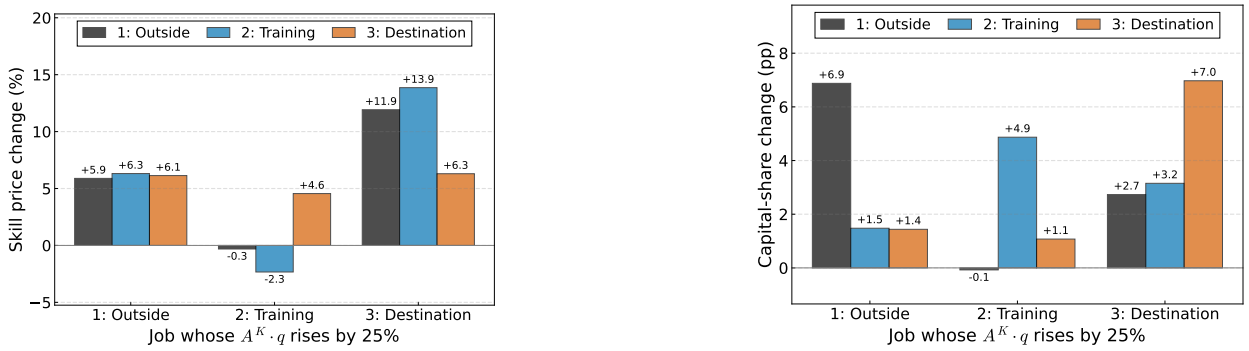
The resulting changes in the skill prices in each occupation are given by:

$$d\log w_1 = \left(\frac{s^K}{s^H}\right)\pi_i\tau - (\pi_3 - \beta\pi_2)\Delta_i,$$
$$d\log w_2 = d\log w_1 - \beta\Delta_i,$$
$$d\log w_3 = d\log w_1 + \Delta_i,$$

where π_i denotes the occupational expenditure share of occupation i in the initial steady state.[30] *Training automation thus generates the largest increase in skill prices in the destination occupation, while destination automation generates the largest increase in skill prices in the training occupation. In contrast, outside automation raises all skill prices by the same proportion. The occupation directly experiencing the automation shock experiences the largest increase in its capital share, while changes in the capital share of the other occupations follow the ranking of changes in skill prices, as implied by Equation (17).*

The intuition for these results is simple. Young workers reallocate across the outside occupation and the career ladder. However, entering the career ladder supplies labor to both the training and destination occupations. Labor reallocation therefore cannot absorb a shift in labor demand from one stage of the career ladder to the other. By contrast, automation outside the ladder can be accommodated by expanding entry into the career path, without changing relative skill prices. Figure 6 illustrates these mechanisms in the nonlinear model with positive idiosyncratic productivity dispersion. The figure shows the impacts of a 25% increase in capital productivity in each occupation and plots the resulting equilibrium changes in skill prices (panel (a)) and capital shares (panel (b)) for all occupations. This stylized economy predicts very stark patterns of occupational mobility. In Sections 5 and 6, I investigate the quantitative relevance of this mechanism in a model calibrated to match observed patterns of occupational mobility.

Figure 6: Cross-occupation spillovers from an automation shock



(a) Changes in skill prices

(b) Changes in capital shares

Notes: Figure shows the nonlinear changes between stationary equilibria following a 25% increase in capital productivity A_i^K, applied to each occupation separately. The three-occupation economy sets $\nu = 0.2$, $\beta = 0.95^{20}$, $s = 0.1$, $\eta = 2$, and $\lambda = 1$. The initial equilibrium is calibrated to $w_1 = 1$, $w_3 = 3$, $e_2^y = 0.5$, and $s_j^K = 0.4$ in every occupation, with the training wage and supporting technology primitives determined accordingly. All other fundamentals remain fixed within each experiment. Groups along the horizontal axis identify the shocked occupation, and colors identify the responding occupation. Skill-price changes are reported in percent and capital-share changes in percentage points.

[30] Under the stated assumptions, occupational expenditure shares in the initial equilibrium are given by:

$$(\pi_1, \pi_2, \pi_3) = \frac{\left(2w_1(1-e_2^y),\ w_2e_2^y,\ w_3e_2^y\right)}{2w_1(1-e_2^y) + (w_2+w_3)e_2^y}.$$

4.3 Exact hat algebra

The previous sections characterized how the stationary equilibrium of this economy responds locally to exogenous technology shocks. These comparative statics revealed that the first-order responses of equilibrium aggregates depend only on occupational expenditure, employment, and wage bill shares in the initial equilibrium and parameters. However, standard exact hat-algebra methods for dynamic discrete choice models, commonly used in the trade and spatial literature, also apply to this economy (Smeets et al. (2025), Traiberman (2019)). In particular, solving for the full nonlinear response of this economy to an exogenous automation shock $\boldsymbol{A}^K \cdot \boldsymbol{q}$ does not require recovering every fundamental of the initial stationary equilibrium. Conditional on values for $\{\nu, \beta, \lambda, \eta\}$, sufficient data consist of (i) occupation-level expenditure and factor-cost shares from the labor demand side and (ii) employment shares, occupational transition probabilities, and average wages by worker type, age, initial occupation (for old workers), and current occupation from the labor supply side. Combined with employment, average wages determine each group's share of its current occupation's wage bill and its contribution to labor supply.[31]

In my data setting, however, I do not observe old workers' wages in each occupation conditional on their initial occupation.[32] In the next section, I therefore describe how I use the structure of the occupational choice model to recover the parameters governing the labor supply model from data on occupational choices and cross-sectional earnings by age and occupation. This procedure allows me to construct estimates of the missing labor supply side variables necessary to perform counterfactuals. Section 6.1 describes how I externally calibrate λ and η and the remaining cost shares that govern the labor demand side of the economy.

5 Mapping model to data: labor supply

The previous section developed a dynamic model of occupational choice with a flexible, reduced-form learning technology. This section describes how I use the structure of the occupational choice model to recover estimates of the parameters governing the behavior of labor supply using data on occupational transitions from Revelio and occupational employment shares and earnings from the ACS. Section 5.1 first describes two model extensions that I introduce when taking the model to the data. Section 5.2 then outlines the identification approach and Section 5.3 provides details on the empirical implementation. Finally, in Section 5.4 I relate the recovered learning technology to observable measures of occupational similarity, while Section 5.5 assesses the empirical relevance of complementarities in labor supply and examines how accounting for on-the-job learning shapes model-implied labor supply elasticities.

5.1 Model extensions

Before taking the model to the data, I add two extensions. First, I introduce M persistent worker types that are defined based on observable characteristics. Types are indexed by $m \in \{1, \ldots, M\}$, associated with a fixed mass μ_m in each cohort. I allow each worker type to have their own initial vector of skills $\{\psi^y_{mj}\}$, on-the-job learning-technology $\{\gamma_{mjk}\}$, and non-wage utility shifters $\{a_{mj}\}$. I incorporate worker heterogeneity into the quantitative model both to study distributional outcomes in response to an automation shock and to distinguish the effects of on-the-job learning from selection into occupations based on workers' initial skills

[31] These are the same objects that appeared in the first-order decompositions.

[32] Although Revelio Labs constructs salary estimates for individual positions using a proprietary salary prediction model, I use these estimates only for benchmarking my results.

and preferences.

Second, I define a working life to be four ten-year periods, which I map to the following age groups in the data: 22-31, 32-41, 42-51, and 52-61. I incorporate additional periods into the model in order to allow for continued occupational reallocation among incumbent cohorts during the transition following an initial automation shock. In the baseline estimation, I assume that workers only accumulate skills on-the-job in the first decade of their working life. This simplification is motivated by the substantial slowdown in wage growth and change in occupational skill content after age 40 that I observe in the data, as well as sample size and selection concerns in restricting the estimation sample to older LinkedIn users for whom I can observe labor market trajectories for over twenty years. Under this assumption, careers are effectively distinguished by an initial learning period (ages 22-31) and a post-learning period (corresponding to ages 32-41, 42-51, and 52-61). In each of the post-learning periods, employment decisions are static and occupational employment shares are identical across age groups, conditional on persistent worker type and the first period occupational choice. In Section 5.3, I discuss how the baseline calibration of the model does at matching occupational employment shares and wages for workers aged 42-51 and 52-61.[33]

5.2 Identification approach

I externally calibrate the discount factor β to 0.95^{10}, based on a common value of 0.95 for an annual discount factor. In this section, I first describe how I recover estimates of type-specific initial skills, bilateral learning, and amenities from observable type-specific moments constructed in an initial steady state, conditional on a value for the shape parameter of the idiosyncratic productivity shocks, $1/\nu$. I then describe how I discipline ν. To provide intuition for the first part of this procedure, I use young workers' average earnings and occupational employment shares to infer each type's initial earnings capacities across occupations. I then use workers' occupational choices and potential earnings in each occupation to identify occupation-specific amenities. Jobs that attract a greater share of workers despite low current wages and worse future career opportunities must offer more favorable amenities. Because amenities are assumed to be age-invariant, they cannot by themselves explain how occupational choices evolve as workers gain experience. Occupational transitions thus identify which paths offer greater learning value.

Type-level model inversion. Let the superscript y ("young") denote the 22-31 age group. Under the distributional assumptions on idiosyncratic productivity shocks, mean log earnings among young workers of type m satisfies the following expression:

$$\mathbb{E}[\log w_j \psi^y_{mj} \varepsilon_j | j \text{ chosen}] = \log(w_j \psi^y_{mj}) + \underbrace{\nu(\gamma_E - (\log e^y_{mj}))}_{\text{Selection adjustment}}, \tag{22}$$

where e^y_{mj} denotes the employment share of young type m workers in occupation j. The first term on the right side of this equation reflects the common initial earnings capacity of type m workers in occupation j, while the second term reflects selection on the basis of idiosyncratic productivity realizations: a rarely chosen occupation requires a more favorable productivity draw. The strength of this selection correction is increasing in the dispersion of idiosyncratic productivity shocks, which is governed by ν. Conditional on a value for ν, I infer $\log(w_j \psi^y_{mj})$ directly from average log wages and the employment share of type m workers in occupation j at ages 22-31. However, the skill price in occupation j cannot be separately

[33] In ongoing work, I am extending the learning period to include the 32-41 age group.

identified from the overall level of initial skills in j. I therefore normalize $w_j = 1$ for each occupation j in the initial equilibrium in the quantitative application in Section 6. This normalization is inconsequential for the recovery of the other parameters in this section and the counterfactuals.

To recover non-wage amenities, recall from Equation (7) that ex ante expected lifetime utility is equalized across young workers within a given type, while expected period utility in the post-learning period (in each of the later three age groups, which I will denote as "old" for simplicity) is equalized across old workers within a type, conditional on first-period occupational choice. Denote the share of type m workers employed in occupation k when old, conditional on having worked in occupation j when young as $e^{o|j}_{mk}$. The share of type m workers employed in occupation k in the post-learning period is then given by $e^{o}_{mk} = \sum_j e^{y}_{mj} \cdot e^{o|j}_{mk}$. Combining these conditions yields the following linear system, which can be used to solve for the J occupation-level amenities for type m:

$$\begin{aligned} \sum_j \left[\frac{e^{y}_{mj} e^{o|j}_{mk}}{e^{o}_{mk}} \Big(\mathbb{E}\big[\log(w_j \psi^{y}_{mj} \varepsilon^{y}_{j}) \mid j \text{ chosen young} \big] + a_{mj} \Big) \right] \\ + B \cdot \Big(\mathbb{E}\big[\log(w_k \psi^{y}_{mk} \gamma_{mjk} \varepsilon^{o}_{k}) \mid k \text{ chosen old} \big] + a_{mk} \Big) = C_m. \end{aligned} \tag{23}$$

where $B \coloneqq \beta + \beta^2 + \beta^3$ and C_m denotes the common expected lifetime utility of young workers of type m. I solve for C_m jointly with occupational amenities, subject to a normalization of the overall amenity level within each type.[34] Intuitively, amenities substitute for wage compensation at both stages of the career. If an occupation offers higher utility when young, through either wages or amenities, then it must offer a lower continuation value. After the initial learning period, a more attractive destination amenity likewise permits lower expected wages. Averaging over all of the initial occupations *observed* for each destination k yields the linear system in (23). Thus, type-specific moments on average log wages by occupation and age, occupational transition shares $e^{o|j}_{mk}$, and initial employment shares e^{y}_{mj}, identify amenities through the occupation's own non-wage rewards and those of its incoming origins. In practice, I measure $\mathbb{E}\big[\log(w_k \psi^{y}_{mk} \gamma_{mjk} \varepsilon^{o}_{k}) \mid k \text{ chosen old} \big]$ as average log wages of type m workers aged 32-41 working in occupation k. Similarly, I measure $e^{o|j}_{mk}$ as the share of workers who worked in occupation j when young and later moved to occupation k ten years later. Details on the construction of these moments are provided in the next section.

Lastly, I infer the level of learning between occupations j and k for type m, γ_{mjk}, from the share of old workers who are employed in occupation k after working in occupation j during the initial learning period, which is given by the following expression in the model:

$$\nu \log e^{o|j}_{mk} = \underbrace{\log \gamma_{mjk}}_{\text{specific to } j \to k} - \underbrace{\frac{\nu \log e^{y}_{mj} - \log(w_j \psi^{y}_{mj}) - a_{mj}}{B}}_{\text{specific to } j} + \underbrace{\log(w_k \psi^{y}_{mk}) + a_{mk}}_{\text{specific to } k} - \underbrace{\frac{C_m - (1+B)\nu\gamma_E}{B}}_{\text{common}}. \tag{24}$$

Inferred learning from j to k is therefore increasing in the probability of transitioning from j to k, conditional on C_m, the young employment share e^{y}_{mj}, and initial earnings capacities and amenities in both occupations. Intuitively, a destination that attracts more workers from a given origin must offer larger learning-related earnings gains, holding fixed the other determinants of occupational choice. The expression in (24) implies

[34] This normalization does not affect the recovered learning estimates.

the following restriction on the pattern of inferred learning:

$$\log\left(\frac{\gamma_{mjk}/\gamma_{mj\ell}}{\gamma_{mj'k}/\gamma_{mj'\ell}}\right) = \nu \log\left(\frac{e^{o|j}_{mk}/e^{o|j}_{m\ell}}{e^{o|j'}_{mk}/e^{o|j'}_{m\ell}}\right). \tag{25}$$

To provide intuition for the implications of this restriction, note that in the job profiles data, working as a financial manager is relatively more common than working as a software manager among workers who first began their careers working as a financial analyst, relative to observably similar users who first began their careers working as a software programmer. Through the lens of the model, on-the-job learning rationalizes such systematic differences in occupational transitions: working as a financial analyst increases financial managerial skill more than software managerial skill, relative to working as a software programmer.

My approach to inferring workers' potential payoffs from occupational choices is related to recent work by Cortes and Gallipoli (2018), who develop an empirical model of occupational switching with broadly defined mobility costs to interpret month-to-month occupational transitions in the Current Population Survey. However, my model imposes additional structure on these payoffs by assuming that on-the-job learning drives systematic patterns of occupational switching over a longer time horizon. The next section discusses the potential limitations of interpreting the recovered learning parameters as productivity gains. I also discuss how the model does at matching within-individual earnings changes across pairs of occupations.

Fréchet shape parameter. As shown in Section 3, the Fréchet shape parameter $1/\nu$ plays an important role in governing occupational labor supply elasticities.[35] A smaller value of ν implies greater responsiveness of labor supply to changes in occupational skill prices. Equation (25) also shows how ν influences the recovered learning estimates. When idiosyncratic productivity shocks are more dispersed, larger differences in learning are required to rationalize observed patterns of occupational transitions.

In the absence of remaining cross-sectional variation in choice moments to identify ν, I follow an existing literature that uses moments of the cross-sectional distribution of wages to discipline this parameter (Caunedo et al. (2023), Cortes and Gallipoli (2018), Hsieh et al. (2019)). In particular, log wages among young workers of a given type employed in occupation j in the model follow a Gumbel distribution with scale parameter ν and mean given by Equation (22). I calibrate ν to match the variance of log wages within type-occupation cells among young workers in the ACS. Additional details of this procedure are provided in Appendix C.3. My baseline calibration sets $\nu = 0.4$, which is close to the value implied by the wage-distribution-based estimate from Burstein et al. (2019) ($\nu = 0.42$), but lower than the value implied by their baseline estimate ($\nu = 0.55$), which exploits exogenous variation in occupational skill prices over time.[36]

5.3 Data and empirical implementation

I assume that the U.S. economy is in an initial steady state during 2010–2022. I begin the estimation period in 2010 to limit the influence of the Great Recession, and I end the estimation period in 2022 to precede the widespread adoption of generative AI following the release of ChatGPT in November 2022. When constructing moments in the ACS, I exclude 2020 to reduce the influence of pandemic-related disruptions.

[35] While the model in Section 3 considered only one worker type, the analogous expression for ε^S_{jk} in the presence of multiple worker types sums over the type-specific responses, weighted by their contribution to j's effective labor supply.

[36] The wage-distribution-based estimate of the Fréchet shape parameter in Burstein et al. (2019) uses an externally calibrated dependence parameter from Hsieh et al. (2019) to introduce correlation in productivity draws across occupations. 0.42 is the implied value without the adjustment.

Defining worker types. I consider eight persistent worker types based on educational characteristics that I observe in both the Revelio data and the ACS, the latter of which I use to construct occupational employment shares and average log earnings. The first five types comprise workers with a bachelor's degree and are distinguished by their undergraduate major, motivated by existing work that shows that bachelor fields of study are informative about workers' comparative advantage and preferences across occupations (Smeets et al. (2025)).[37] I construct these categories by clustering majors according to their occupational employment distributions among workers aged 22–31 in the ACS. The resulting groups are: Engineering, Computing and Mathematics; Health and Life Sciences; Business, Economics and Government; Care and Education; and Media, Letters and Public Service.[38] Additional details are provided in Appendix C.1, where I also show how occupational employment shares and average log earnings differ across the five bachelor types. The three remaining worker types are high school graduates without a college degree, associate degree holders without a bachelor's degree, and workers without a high school diploma.[39]

As discussed in Section 2, I identify age in the job profiles data using the graduation dates listed on users' profiles. However, I cannot reliably identify LinkedIn users without a high school diploma, since users without a degree listed on their profile may have simply omitted their educational history. These workers are also unlikely to be active on LinkedIn. I therefore assume that workers without a high school diploma share the same on-the-job learning technology as workers whose highest qualification is a high school diploma in the baseline model. I discuss below how the model does at matching average earnings and occupational employment shares for this (relatively small) group of workers in the ACS.[40] Finally, I set the mass of each worker type μ_m equal to its employment share among workers aged 22–31 in the ACS. Figure A.10 compares the distribution of worker types in the ACS and Revelio. Although Revelio over-represents workers with bachelor's degrees relative to the general population, I show below that the type-specific occupational transition networks that I construct from the Revelio data closely predict the occupational employment distributions for workers without a bachelor's degree in the ACS. I also find that annual transition rates across broad occupational categories closely track those observed in the Current Population Survey (CPS).

Occupations. I begin with the occupational classification in the 2018 ACS, which is already an aggregation of the detailed codes in the 2018 SOC taxonomy (525 occupations, down from 848 occupations, excluding military occupations). I further aggregate occupations to improve coverage across worker types and age groups in the ACS, while preserving distinctions among professional and managerial occupations where feasible. The resulting classification contains 64 occupations. Table A.2 reports the classification along with details of the procedure used to combine occupations. Compared with using the 95 non-military SOC minor groups directly, this classification preserves greater detail among several professional and managerial occupations (SOC major categories 11-29), while consolidating service and production occupations that are sparsely represented in Revelio and among college-educated workers. As shown in Appendix A.3.2, these categories also have a low level of exposure to generative AI-driven automation.[41]

[37] Using administrative data on applications to fields of study from Denmark, Smeets et al. (2025) estimate a structural model of field choice that allows for latent heterogeneity across workers. I abstract from educational choices as an additional margin of adjustment and focus my analysis on on-the-job skill accumulation decisions.

[38] Graduate-degree holders remain in their corresponding undergraduate-major group.

[39] Data limitations preclude further disaggregation within these groups.

[40] Workers without a high school degree comprise less than 6% of young workers in the pooled ACS data.

[41] Some professional occupations, such as physicians and lawyers, require educational qualifications beyond those held by certain worker types. I treat the corresponding type–occupation cells as restricted in the model.

Employment shares and wages. I construct employment shares and average log wages for each type and age group using pooled ACS data from 2010-2022, excluding 2020. Wages are annual earnings divided by annual hours, where earnings are deflated to 2012 dollars using the Personal Consumption Expenditures Price Index. I winsorize the hourly wage distribution at 0.1% and 99.9% percentiles. Appendix C.1 provides additional details on my ACS sample and data construction.

Occupational transition networks. To construct type-specific occupational transition networks corresponding to young workers in the initial learning period $e^{o|j}_{mk}$, I focus on users aged 22–31 in 2010-2012 in the Revelio data, using the measure of age defined in Section 2. I assign types to users by their highest level of education by 2010. For the worker types without a bachelor's degree, I restrict to users that do not eventually report higher degree attainment on their profile to align the sample more closely with the ACS.[42] I record users' initial occupation during 2010-2012 and their ultimate occupation ten years later, in 2020–2022, when they are ages 32–41. Transitions are thus defined by the users' endpoint occupations and may involve intermediate switches, as in the empirical analysis of Section 2.

To address sampling zeros and imprecisely measured transitions, I smooth the empirical transition networks using an empirical Bayes procedure that pools information across origin occupations within each worker type. This procedure stabilizes sparsely observed transitions while keeping well-supported rows close to their empirical frequencies.In practice, the large sample size of the Revelio data ensures that almost all origin occupations are well-supported for the college-educated (bachelor's and associate's) worker groups, limiting the influence of the priors on their transition networks. For the high-school only group, smoothing plays a somewhat more important role, given that it is the smallest population in the Revelio data. Appendix C.2 provides details on the smoothing procedure and reports coverage statistics for each worker type.

Applying the resulting 10-year transition networks constructed from the Revelio data to the ACS occupational employment shares of workers aged 22–31 yields predicted occupational employment shares for workers aged 32–41 that closely align with their ACS counterparts. Within-type correlations across occupations are above 0.90 for all types, except "Media, letters, and public service" majors (0.73) and "Business, Economics, and Government" majors (0.78). Notably, for workers without a college degree, the correlation is above 0.95. In a final step, I further adjust the transition networks using Iterative Proportional Fitting (IPF), so that, when composed with the ACS young employment shares, the transition networks exactly reproduce the ACS occupational employment shares of workers aged 32–41 within each type.[43] The adjustment procedure is described in Appendix C.2, and explains that, in practice, it amounts to a correction to destination shares that accounts for type-specific differences in occupational coverage in Revelio relative to the ACS.

Internal data validation. By construction, the model-based inversion procedure implies that the model exactly matches occupational employment shares and average log earnings within each worker type for the 22–31 and 32–41 age groups in the pooled 2010–2022 ACS data.[44] I perform three robustness checks to address potential data alignment issues and over-fitting concerns.

[42] In the ACS, I only observe educational attainment as of the survey date. I exclude users who were enrolled in a form of schooling in the last three months that is higher than their current highest degree qualification for the non-bachelor types.

[43] Iterative Proportional Fitting (IPF) is a procedure that alternately rescales the rows and columns of a matrix to match specified marginal distributions. This adjustment does not apply to the users without a high school degree, as I do not directly construct a transition network for them.

[44] The procedure does not impose that the model-implied moments for workers without a high school degree aged 32–41 match their empirical counterparts. The correlation for average log earnings is 0.72 and 0.99 for employment shares.

First, I verify that the implied learning estimates are similar when I instead measure mean log earnings for workers aged 22–31 in the 2011–2012 data and workers aged 32–41 in the 2021–2022 data. This alternative approach follows the same set of cohorts across the two age groups, providing a cohort-aligned measure of within-occupation earnings profiles that maps more closely to the cohorts that I use to construct the transition networks. I use the pooled ACS estimates in baseline estimation to reduce noise in the estimates.

Second, since the identification procedure uses data only on the first two age groups, I check how the model-implied occupational employment shares and average log earnings for each type compare to the data for the last two age groups (workers ages 42-51 and 52-61), with the caveat that these comparisons mix differences in cohort composition with age since these moments are constructed from the pooled 2010-2022 ACS data. The baseline model assumes no learning after the first period, and so these moments are mechanically flat after the initial learning period in the model. As expected, the model under-predicts average log wages for the two older groups, missing about a quarter of the difference in average log wages between the youngest and oldest age groups on average. However, it matches the correlation of average log wages across occupations well.[45] The model also matches the occupational employment shares of the older two age groups very well. The correlation is above 0.95 for all types.

Third, to assess potential overfitting concerns from matching bilateral occupational flows (cf. Dingel and Tintelnot (2026)), I evaluate the out-of-sample predictive performance of the estimated transition networks. In particular, I randomly split workers in my Revelio sample into two halves and re-estimate the networks on both halves. I then assess their predictions for movers in the other half, compared to a benchmark that assigns destinations according to type-specific destination shares, computed across all movers. Repeating this procedure 100 times (resampling users across iterations) shows that the origin-specific patterns that I recover in the transition networks predict mobility well beyond destinations' overall popularity for all types. I also assess how the constructed transition networks perform out-of-sample relative to a gravity-like regression that includes detailed measures of occupational similarity, in addition to origin and destination fixed effects. As shown in Figure A.11, the empirical transition networks that I construct outperform the gravity benchmark for all types, suggesting that they contain stable predictive information beyond detailed measures of occupational similarity.[46]

Results and validation. Figure 7 plots the recovered (log) initial skill level for each worker type in each occupation, relative to the type-weighted average skill level in each occupation. The figure shows that initial skills differ across types within an occupation in a predictable manner. For example, "Business, Economics, and government" majors tend to earn more in business and financial occupations, while "Engineering, computing, and math" majors tend to earn more in computer, math, and engineering occupations. Appendix C.4 reports heatmaps that show the recovered estimates of type-specific amenities and learning. Figure A.14 shows the ten (directed) occupational pairs associated with the highest learning estimates for each bachelor's type, and Figure A.13 shows the ten (directed) occupational pairs associated with the highest learning estimates for the two non-bachelor worker types. In general, I find meaningful differences in inferred learning across types, as well as across occupations within types. In Section 5.4 below, I analyze how the recovered learning technology relates to observable measures of occupational similarity.

An important caveat to my results is that the recovered learning parameters may capture remaining unob-

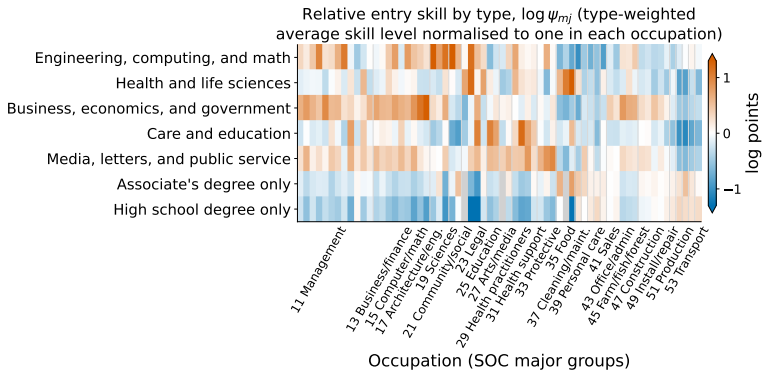
[45] The correlation is above 0.95 for all types, except for the "less than high school" type for whom it is around 0.7.

[46] I estimate the gravity model with a PPML regression that includes movers only. The included covariates are described in Section 5.4, where I show that these covariates do have meaningful predictive value for transitions (and inferred learning).

served heterogeneity within worker types, as well as unmodeled mobility frictions. As a preliminary check of the model's implications for wage growth, I compare its predictions with within-user changes in Revelio's salary predictions, which are not used in the model inversion. Revelio predicts annual salaries using a model trained on salary information from job postings, visa applications, and self-reported compensation. The resulting salary-growth measure reflects changes in predicted pay that are based on observable job characteristics, rather than individual histories.

For each worker type and origin–destination occupation pair, I compute the average ten-year change in users' predicted log salaries and compare it with the model's expected log wage growth conditional on the same occupational path. The model prediction thus accounts for initial earnings capacities, accumulated skills, and selection on idiosyncratic productivity in both periods. Across occupational moves, model-implied wage growth is strongly correlated with the average changes in Revelio's salary predictions, with correlations ranging from 0.68 to 0.77 across worker types. However, the slope from regressing Revelio salary growth on model-implied wage growth is below one for all types (ranging from 0.55 for the engineering type to 0.77 for high-school degree holders), indicating that differences in model-implied wage growth correspond to smaller differences in salary growth computed from Revelio's position-level salary predictions.[47]

Figure 7: Inferred skills across types and occupations



Notes: Figure shows the log of the type-specific initial earnings capacities relative to the type-weighted average mean. Gray bars denote occupations that are closed to a given type due to educational requirements.

5.4 Economic content of the recovered learning technology

To provide insight into the economic content of the recovered learning technology, I investigate how inferred learning relates to observable measures of occupational similarity. More specifically, I estimate weighted regressions of the following form, separately for each worker type:

$$\log \gamma_{mj\ell} = \alpha^{O}_{mj} + \alpha^{D}_{m\ell} + \beta_{m,K} \cdot \text{Distance}^{\text{knowledge}}_{j\ell} + \beta_{m,S} \cdot \text{Distance}^{\text{skill}}_{j\ell} \tag{26}$$

$$+ \beta_{m,T} \cdot \text{Distance}^{\text{task}}_{j\ell} + \beta_{m,I} \cdot \text{Distance}^{\text{industry}}_{j\ell} + \beta_{m,G} \cdot G_{j\ell} + \varepsilon_{mj\ell}. \tag{27}$$

[47] The model also over-predicts average growth among occupational stayers. Most of this discrepancy is concentrated among users in the same recorded position at both endpoints, for whom predicted salary growth depends on Revelio's within-position salary adjustments. Additionally, the Fréchet specification implies that expected log wage growth is additive in origin and destination components within each worker type. This exercise therefore assesses the model's joint earnings implications, rather than independently validating bilateral learning patterns.

Each observation is a directed pair of occupations. Origin fixed effects (α^O_{mj}) absorb an origin's general learning opportunities, while destination fixed effects ($\alpha^D_{m\ell}$) absorb the average learning gains associated with each destination.[48] $\text{Distance}^{\text{knowledge}}_{j\ell}$, $\text{Distance}^{\text{skill}}_{j\ell}$, and $\text{Distance}^{\text{task}}_{j\ell}$ are measures of distances between j and ℓ's knowledge requirements, skill requirements, and task content, respectively, which I construct from O*NET. In O*NET, knowledge refers to the subject-matter expertise required in an occupation, such as economics, engineering, or medicine, while skills describe the capabilities used to perform work activities, such as critical thinking, coordination, or troubleshooting. In contrast, tasks refer to the work activities that individuals actually perform on the job, which also underpin the automation exposure measures described in Section 2. Similarly, $\text{Distance}^{\text{industry}}_{j\ell}$ measures the distance between j and ℓ's employment distributions across 20 NAICS sectors, which I construct using the 2022 OEWS.[49] Finally, $G_{j\ell}$ is an indicator for membership in the same SOC major (2-digit) group. For ease of interpretation, I scale the continuous distance measures by their standard deviation across all pairs after removing origin and destination fixed effects, weighting pairs by the mass of movers.

The left panel of Figure 8 reports the estimated coefficients from separate type-specific regressions that include only one covariate at a time alongside the fixed effects, while the right panel reports the within R^2 for each of these regressions. Across all worker types, greater distance in knowledge, skills, task content, or industry composition is associated with lower inferred learning. For example, a one-standard-deviation reduction in residualized knowledge distance is associated with 0.19–0.25 log points higher inferred learning across types, translating to a 21–28% higher learning multiplier $\gamma_{mj\ell}$. Table A.4 reports the complete table for the regression specification including all covariates jointly. In general, similarity in knowledge requirements and industry composition provides the greatest explanatory power of recovered learning on average. Taken together, these statistical decompositions of the recovered learning technology are consistent with an important role for the accumulation of domain-specific human capital on-the-job. More broadly, these results reveal that measures of occupational similarity explain a meaningful share of the bilateral variation in longer-run patterns of occupational mobility, suggesting that gravity-model predictions could provide an informative center for the empirical Bayes prior used in the construction of empirical occupational transition networks. In my particular data setting, this procedure yields very similar transition networks (see Figure A.11), consistent with the limited influence of regularization on rows with substantial support.

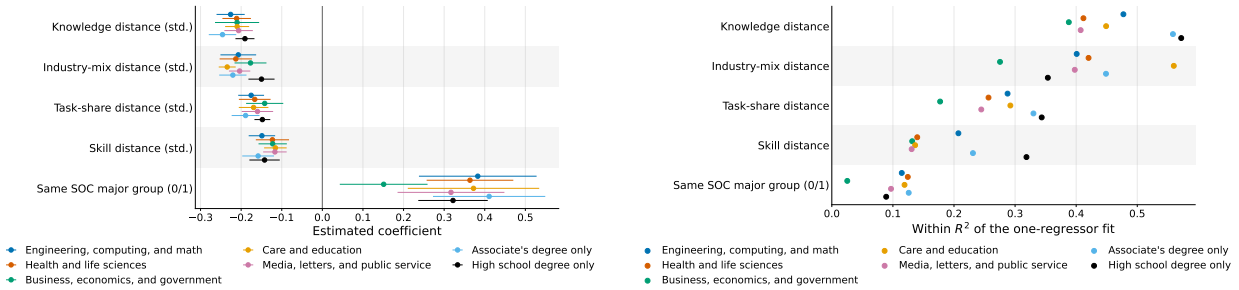
Finally, Figure A.16 shows that inferred learning is also *directional*. In particular, for each pair of occupations, I compare learning in one direction with that in the reverse direction. I find that this difference tends to be larger for moves towards higher-paying occupations and occupations involving more reasoning, decision-making, and information processing, as measured by O*NET. Thus, occupational similarity alone does not fully characterize the recovered learning technology. This asymmetry is consistent with the idea that experience acquired in earlier jobs prepares workers for more complex tasks and greater responsibility.[50]

[48] Under the choice inversion in Equation (24), the residual from projecting $\log \gamma_{mj\ell}$ on origin and destination fixed effects equals ν times the residual from the same weighted projection of the log conditional transition probability. Thus, this exercise characterizes the occupational-content patterns in mobility that the model interprets as bilateral learning.

[49] Knowledge, skill, and task distances are root-mean-square differences between occupations' detailed O*NET elements, with each element standardized across occupations. Industry-mix distance is half the sum of absolute differences in occupations' industry employment shares, which measures total-variation distance. Results are similar when using alternative distance measures (e.g., cosine distance) and aggregations of O*NET elements.

[50] For the engineering type, the strongest directional patterns instead involve moves towards occupations emphasizing interpersonal skills, autonomy, and decision-making, consistent with progression from technical work towards roles involving coordination and management.

Figure 8: Relationship between inferred learning and observable measures of occupational similarity



(a) Estimated coefficients

(b) Within-R^2

Notes: Figure reports results from separate type-specific regressions of Equation (26), each including one distance measure alongside origin and destination fixed effects. Pairs are weighted by the mass of movers, and distance measures are standardized across pairs. Panel (a) shows estimated coefficients and 95% confidence intervals. Standard errors are two-way clustered by origin and destination occupation but do not account for estimation uncertainty in recovered learning. Panel (b) reports the corresponding within-R^2, measuring the share of variation in log learning explained by each distance measure after removing the origin and destination fixed effects. Results are similar when restricting to directed pairs with at least 30 movers and in unweighted regressions. Appendix C.3 reports the results for the regression specification including all covariates jointly.

5.5 Complementarities in labor supply

Finally, I use the recovered skill and learning estimates to compute model-implied own- and cross-wage elasticities of occupation-level labor supply, using a modified version of Equation (18), adapted for multiple worker types and age groups. Figure 9 plots the distributions of model-implied labor supply elasticities with those implied by a static model of occupational choice calibrated to directly match observed employment shares and mean log wages by worker type, age group, and occupation in the ACS. This static comparison retains observed differences in occupational sorting and earnings across types and ages but abstracts from forward-looking career choices and history-dependent skill accumulation. This static model is thus closer to approaches in existing work that study worker heterogeneity along observable dimensions that are available in cross-sectional data, such as age and education (Acemoglu and Restrepo (2022), Burstein et al. (2019), Caunedo et al. (2023)).[51]

The left panel of Figure 9 shows that positive cross-wage labor supply elasticities are relatively rare in practice: on average, the model infers that occupations are labor supply substitutes.[52] Notably, positive cross-wage elasticities are entirely absent from the static model. Moreover, the cross-wage labor supply elasticities in the static model are significantly smaller and less dispersed than those in the dynamic model. The right panel plots the distribution of own-wage labor supply elasticities and reveals a similar difference in magnitudes across the two models. In general, the larger labor supply elasticities in the learning model reflect the importance of endogenous career choices and skill accumulation for accurately capturing long-run flexibility in labor supply, which I elaborate on below.

Figure 10 reports the ten directed occupation pairs with the largest positive and negative cross-wage labor supply elasticities in the learning model. The left panel shows that the strongest complementarities

[51]The static comparison uses the same type composition and equal-sized age groups as the baseline model. Occupational employment and wage profiles at ages 32–41 are assigned to all three post-learning periods, consistent with the baseline's frozen-skills assumption. Results are similar when using the observed profiles for the later age groups and allowing both type composition and cohort sizes to vary across age groups as in the ACS.

[52]Note that the labor supply elasticities within each occupation j must sum to zero (with log utility, labor supply does not respond to uniform wage movements).

connect occupations along recognizable career pathways. For example, effective labor supply in computer and information systems management increases strongly with the skill price of computer occupations. The right panel shows that the strongest negative cross-wage elasticities also correspond to intuitive labor-supply substitutes. For example, effective labor supply in computer occupations falls when skill prices for engineers and architects rise, consistent with these occupations drawing on overlapping pools of workers. Figure 10 further decomposes these elasticities into the three terms in Equation (18). The decomposition reveals that the dynamic composition effect is the dominant source of large cross-wage elasticities in the learning model, both positive and negative. These large negative elasticities highlight an important source of long-run labor-supply flexibility. As relative skill prices change, adjustments in where young workers choose to acquire their skills strengthen substitution patterns in labor supply across occupations.

Figure 11 plots young workers' cross-wage employment elasticities in the decomposition of Equation (18). The figure shows that young worker labor supply also exhibits meaningful complementarities. The ten largest positive cross-wage employment elasticities, computed using Equation (20) and aggregated across types, generally connect occupations along recognizable career pathways. For example, young worker employment in legal support occupations increases with the skill price for lawyers, reflecting the greater value of entering a career pathway that leads to legal work. Importantly, this type of complementarity is also entirely absent from the static model. The overall contribution of the young labor supply response, which captures all complementarities arising purely from forward-looking behavior, to the occupation-level labor supply elasticities is relatively small. This is partly because only young workers sort for learning opportunities in the baseline model and they account for a relatively small share of effective labor.[53]

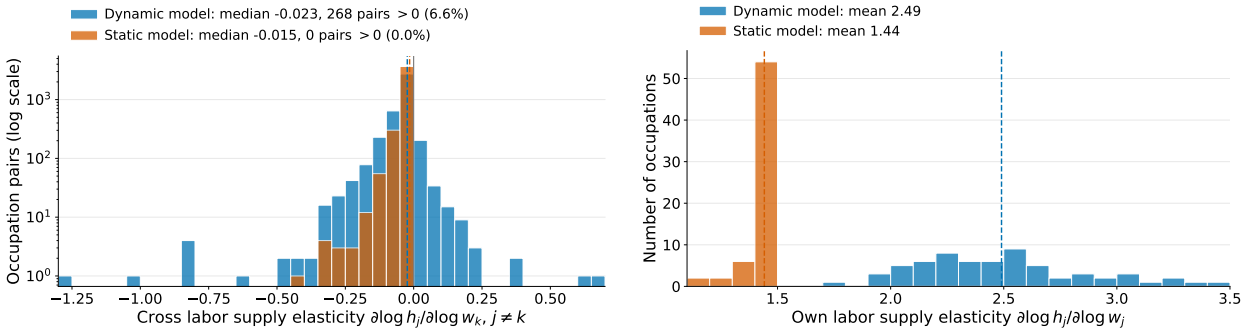
Finally, Figure 12 presents heatmaps of the model-implied cross-wage elasticities in the learning and static models. Recall that with ex ante homogeneous workers and no skill accumulation, the static occupational choice model implies constant columns, excluding the diagonal: for $j \neq k$, $\varepsilon^S_{jk} = -(1-\nu)e_k/\nu$. This expression follows directly from Equation (19), after accounting for the productivity effects of the Fréchet selection model.[54] In the static model with heterogeneous skill types and ages, however, differences in occupational sorting patterns across groups can generate variation across rows: occupations are stronger labor-supply substitutes when they draw on overlapping groups of workers.[55] The dynamic model with on-the-job learning then introduces additional pair-specific variation through both forward-looking occupational choices (the young reallocation term) and the endogenous effects of these choices on the distribution of old worker skills (the dynamic composition term). Figure 12 shows that accounting for these dynamic forces simultaneously strengthens labor-supply substitution across many pairs of occupations while generating complementarities between others.

[53] Young workers are aged 22–31 and constitute one-fourth of the population; the remaining three age groups are post-learning workers. The aggregation weights, however, reflect each group's contribution to effective labor supply.

[54] In a homogeneous worker benchmark, cross-wage elasticities are small when the occupation whose skill price changes accounts for a small employment share.

[55] Fan (2025) and Grigsby (2025) are examples of recent studies that show how skill heterogeneity and transferability across occupations shape cross-occupation labor supply elasticities.

Figure 9: Distribution of model-implied labor supply elasticities

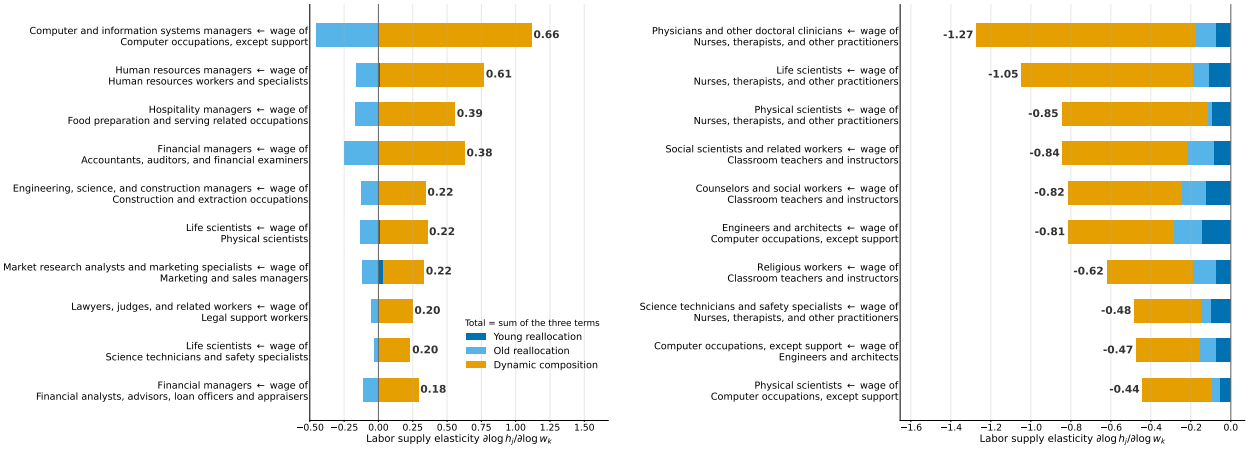


(a) Cross-wage labor supply elasticities

(b) Own-wage labor supply elasticities

Notes: Figure shows the distributions of model-implied cross-wage and own-wage labor supply elasticities, constructed following equation (18), adjusted for multiple worker types and age groups. Static model refers to a model with skill/age heterogeneity only. This computation uses observed data by age and type, with the same value of ν used in the learning model.

Figure 10: Largest positive and negative cross-wage labor supply elasticities in the model

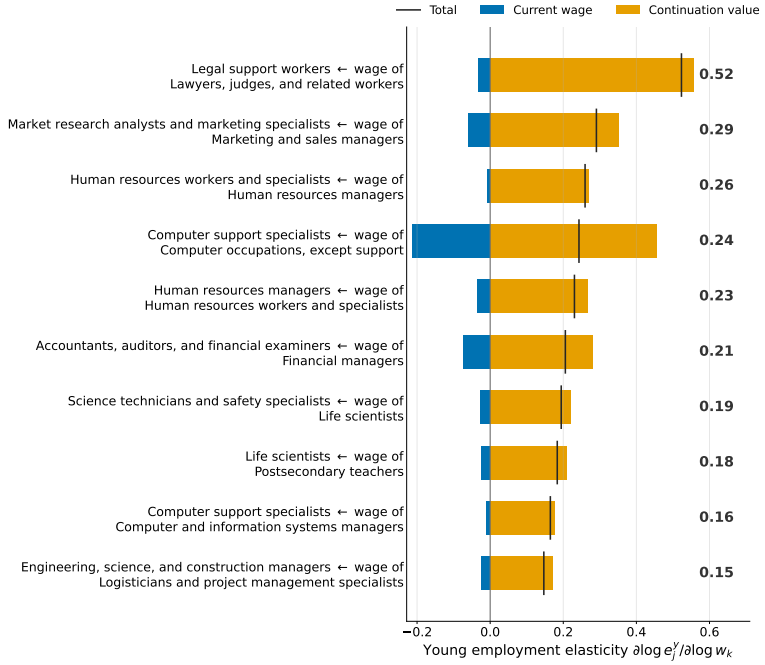


(a) Largest positive elasticities

(b) Largest negative elasticities

Notes: Panels (a) and (b) show the ten (directed) occupational pairs with the largest positive and largest negative cross-wage labor supply elasticities, respectively, constructed following (18), adjusted for multiple worker types and age groups. The young reallocation term is the contribution of (20), the old reallocation term is the contribution of (19), and the dynamic composition effect is the contribution of (21).

Figure 11: Largest positive cross-wage elasticities in young worker labor supply



Notes: Figure shows the ten (directed) occupational pairs with the largest positive cross-wage *young* labor supply elasticities, constructed following Equation (20), adjusted for multiple worker types and age groups.

Figure 12: Heatmaps of model-implied cross-wage labor supply elasticities



(a) Dynamic model with on-the-job learning

(b) Static model with age and skill heterogeneity only

Notes: Figures show the heatmaps of model-implied cross-wage elasticities, constructed following Equation (18), adjusted for multiple types and age groups. Diagonal elements (own wage-elasticities) are not colored in; they are reported separately in Appendix C.4. Panel (a) shows the elasticities in the baseline model, while panel (b) shows the implied elasticities in a model with skill/age heterogeneity only. This computation uses observed data by age and type, with the same calibrated value of ν.

6 Quantitative experiments

In this section, I conduct two types of quantitative experiments aimed at studying how on-the-job learning shapes the individual and aggregate impacts of automation shocks within and across career paths. I first describe the external calibration of the labor-demand parameters and baseline expenditure shares in Section 6.1. Section 6.2 then examines how an improvement in the productivity of the automation technology in a single occupation within an existing career ladder shapes the resulting pattern of automation across occupations. This experiment illustrates the novel mechanisms of the model by which automation shocks propagate across occupations within career paths. Analyzing the transitional dynamics of the economy, I show how accounting for on-the-job learning shapes the incidence of the automation shock on different types and cohorts of workers. Finally, Section 6.3 assesses the distributional consequences of widespread AI-driven improvements in automation technologies, using the occupation-level exposure measures constructed in Section 2 to discipline the pattern of automation shocks across occupations. At the end of the section, I discuss some of the limitations of this forward-looking analysis, several of which I plan to address in future work.

6.1 Calibrating labor demand

As discussed in Section 4.3, it is not necessary to specify the values of $\{A_j^K, q_j, A_j^H\}_j$ in the initial steady state in order to solve for the economy's response to changes in technology, given estimates of the labor-supply side of the model and values for λ, η, and $\{s_j^K\}$. Below, I discuss how I externally calibrate λ and η to existing values from the literature and construct measures of $\{s_j^K\}$ following existing work.

Calibrating λ and η. There is not an established estimate of the elasticity of substitution across occupations. In my baseline calibration, I follow Dvorkin and Monge-Naranjo (2019) and set this elasticity, λ, to one. This value is close to the within-industry estimate of 0.9 reported by Goos et al. (2014), but below the empirical estimates of 1.34 and 1.81 from Caunedo et al. (2023) and Burstein et al. (2019), respectively. I set the elasticity of substitution between capital and labor, η, to two, which is at the upper end of the occupation-level point estimates in Caunedo et al. (2023) and close to the mean occupation-level estimate of Dvorkin and Monge-Naranjo (2019). The baseline calibration thus satisfies the condition that $\eta > \max\{1, \lambda\}$.

While conventional estimates of the elasticity of substitution between capital and labor are often close to or below one, the relevant parameter here governs long-run substitution patterns *within* occupations between workers and the types of capital that are capable of potentially performing their tasks.[56] As discussed in Section 4, the condition that $\eta > \lambda$ implies that alternative ways of performing an occupation's work are more substitutable than the outputs of different occupations, which ensures that improvements in $A_j^K \cdot q_j$ reduce occupational labor demand, all else equal. Similarly, the condition that $\eta > 1$ implies that these improvements will raise an occupation's capital share, all else equal. Conditional on initial values for occupation-level capital shares, a larger gap between η and λ strengthens the effects of automation on relative skill prices that I document below, while a larger value of η will make capital shares more responsive to changes in relative factor costs, as shown in Equations (16) and (17).

Occupation-level capital shares. Similar to recent work by Freund and Mann (2026), I construct occupation-level capital shares using occupational employment and earnings by industry from the OEWS

[56] Hawkins et al. (2015) estimate a plant-level production model that distinguishes between short- and long-run elasticities of substitution between capital and labor. In their specification with capital-augmenting productivity, the estimated elasticity is 1.4 in the short run and 5.6 in the long run.

and industry-level payments to labor and capital from the BEA-BLS Integrated Industry-Level Production Accounts. As a brief overview of this procedure, which is described in detail in Appendix D.1, I allocate industry-level labor and capital income across occupations in proportion to their shares of that industry's wage bill in the OEWS.[57] I then compute occupation-level capital shares from the resulting income totals. Note that the calibrated labor supply model supplies occupation-level wage-bills in the initial steady state. Occupation-level expenditure shares are then implied by the labor shares that I construct from the data and the model-implied wage-bill shares.[58]

6.2 Single automation shock within a common career path

I first examine the long-run incidence of an occupation-specific automation shock across occupations and worker types. For this exercise, I consider an exogenous 22% reduction in the price of capital in computer occupations, the occupational category which includes software developers and programmers but excludes computer support occupations.[59] I then compare long-run outcomes in the model with on-the-job learning against a static benchmark economy without skill accumulation in order to isolate how on-the-job learning matters for occupation- and worker-level outcomes. After comparing the long-run incidence of the automation shock across the two models, I study the transitional dynamics of the model with skill accumulation.

My focus on computer occupations is motivated by the relatively high exposure of this occupation to AI-driven automation (41% of tasks rated as automatable, relative to an employment-weighted mean of 27%), in addition to the fact that it is a quantitatively important labor supply complement to computer and information systems (CIS) managers (whose exposure sits exactly at the employment-weighted mean), as shown in Figure 10. This pairing thus provides a salient example of an uneven automation shock along a career path. The initial steady state of the learning economy reveals the career pathway linking these two occupations. In particular, Figure A.17 shows that the share of young workers in computer occupations is close to the economy-wide average, whereas the corresponding share among CIS managers is less than half as large. Figure A.18 further shows that a majority of CIS managers first work in computer occupations when young. However, this career pathway also coexists with substantial occupational persistence: young workers in both occupations have a high probability of remaining in the same occupation. Appendix D.2 reports the learning, amenity, and initial earnings capacity estimates for both occupations that rationalize these patterns. Importantly, Figure A.19 shows that recovered initial earnings capacities are higher in computer occupations than in CIS management for every worker type.

Static benchmark. To isolate the joint role of forward-looking career choices and endogenous skill accumulation, I construct a static benchmark economy in which workers retain the occupation-specific skills they hold in the learning economy's initial steady state, including skills acquired through past occupational experience. The benchmark economy assigns the same population share to each observed skill portfolio, but holds these portfolios and their population shares fixed in the counterfactual. Workers can still change occupations in response to wages, but their choices no longer affect their future skills and so their occupational choices are static. Labor-demand fundamentals are kept constant and non-wage amenities in the bench-

[57] Appendix D.1 describes the categories of capital included in this allocation.

[58] The model-supplied wage bill shares do not necessarily equal the wage bill shares in the OEWS because they are calibrated to match their age-adjusted analogues in the ACS and do not include variation in hours. Nonetheless, they line up closely.

[59] 22% comes from the mapping from exposure scores to improvements in capital efficiency that I implement in Section 6.3. The results are similar when considering a joint shock to computer occupations, as well as computer support occupations. I focus on a single shock to computer occupations in order to simplify the exposition of the results.

mark are calibrated to reproduce the learning economy's initial occupational allocation and effective labor supply.[60,61] Comparing occupation-level outcomes across the models to the same automation shock thus isolates the contribution of the two dynamic margins present in the baseline model with on-the-job learning. In the learning economy, young workers respond to changes in future returns, and their occupational choices shape their human capital accumulation. In the static benchmark, workers reallocate across occupations, but these dynamic margins are absent. Results are also similar under the alternative static benchmark that I calibrated directly to match observed shares and wages by type and age in Section 5.5.

(Long-run) Occupation-level incidence. Figure 13 shows the responses of employment, skill prices, and capital shares in computer occupations and computer and information systems management in the two economies. Focusing first on the baseline model, the figure shows that CIS managers experience larger increases in skill prices and capital shares relative to the employment-weighted average of all other occupations in the economy.[62] Revisiting the first-order decomposition of skill price responses in Equation (16) helps explain this finding.[63] The components of the first-order decomposition are shown in Figure 14. Through a standard substitution channel, lower skill prices in computer occupations release labor to other occupations, putting downward pressure on skill prices. Conversely, the positive aggregate productivity effect raises skill prices in all occupations, drawing workers away from any given occupation and putting upward pressure on its skill price. On balance, cross-occupation labor supply spillovers raise wages in all occupations, as shown in the rightmost figure of panel (a) of Figure 14. However, the positive contribution of labor supply spillovers is particularly large for CIS managers precisely because of the dynamic composition effect: the decline in the relative skill price for computer occupations *lowers* labor supply in CIS management.

To further clarify the importance of initial occupational choices, Figure 15 shows the changes in employment for young and old workers in the two occupations. Following the initial automation shock, employment in computer occupations falls similarly among both age groups.[64] In contrast, employment in computer management rises, but much more among young workers: direct entry into management increases by 2.06%, compared with only a 0.41% increase in employment among older workers. Because older workers account for most managerial employment, the overall increase in employment is mechanically limited. However, the muted employment response among older workers is not mechanical. It reflects the weakening of the early-career pathway from computer occupations into computer management.

Finally, comparing the long-run responses in the learning economy and the static benchmark in Figure 13, the figure shows that the static benchmark predicts substantially smaller increases in both the skill price and capital share of computer management. This difference reflects how the two models treat the labor-supply relationship between computer occupations and CIS management. In the static model, these occupations

[60] For workers with entry-level skill portfolios, this requires introducing fixed occupation-specific utility shifters that absorb the continuation values associated with each initial occupation in the learning economy's initial equilibrium. These shifters and the distribution of skill portfolios remain fixed in the counterfactual. Results are similar under an alternative no-learning benchmark whose initial equilibrium is recomputed, holding all other fundamentals fixed.

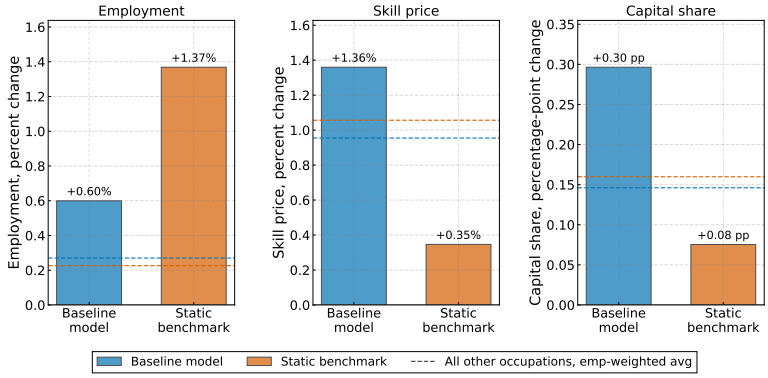
[61] Appendix B.3 shows that this matching implies that the first-order response of the aggregate capital share in the two models would be equivalent if the extent of automation were initially common across all occupations. While the first-order response of the aggregate capital share can in principle differ with heterogeneous initial capital shares, in my current calibration, the (nonlinear) response of the aggregate capital share is very similar in the two economies.

[62] Figure A.24 shows that the response for CIS managers is larger than that for all other occupations in the economy.

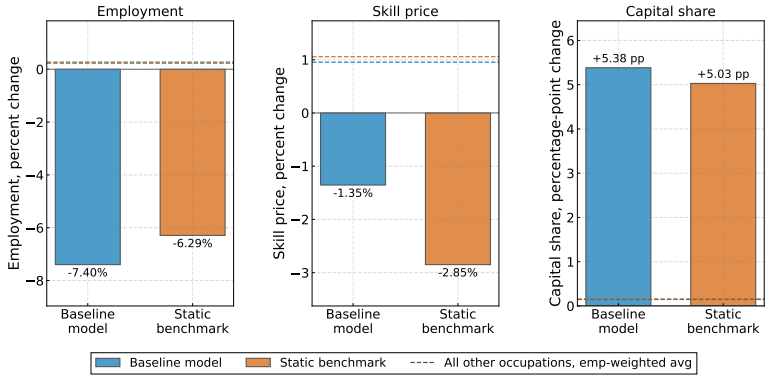
[63] In Figure A.22, occupation-level responses computed using the first-order approximation for the single occupation automation shock closely track those in the nonlinear model.

[64] The slightly larger response among young workers is consistent with the persistence of computer careers: workers entering computer occupations are likely to remain in computer work, so the decline in relative computer skill prices weakens both the current and expected future returns to entry.

Figure 13: Effects of the automation shock on computer occupations and managers



(a) Computer and information systems managers



(b) Computer occupations

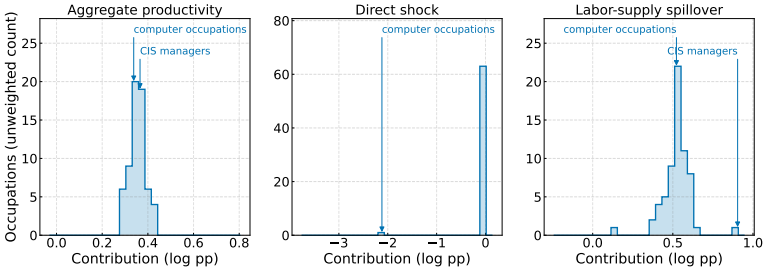
Notes: Figure compares the responses of employment, occupational skill prices, and capital shares to the same reduction in the price of capital used in computer occupations. Panel (a) reports responses for computer and information systems managers; panel (b) reports responses for computer occupations. Dashed lines show the initial-employment-weighted average response of other occupations in each economy, excluding both computer occupations and computer managers.

are strong *substitutes*: declining computer skill prices encourage the overlapping set of workers that are skilled in both computer occupations and management to reallocate toward management. As illustrated in the rightmost figure of panel (b) of Figure 14, this cross-occupation labor supply response significantly dampens the increase in managerial skill prices in the static model.[65] In contrast, in the dynamic model with learning, these two occupations are labor supply *complements* and the negative labor demand shock to computer occupations effectively acts as a negative labor supply shock to computer managerial work.

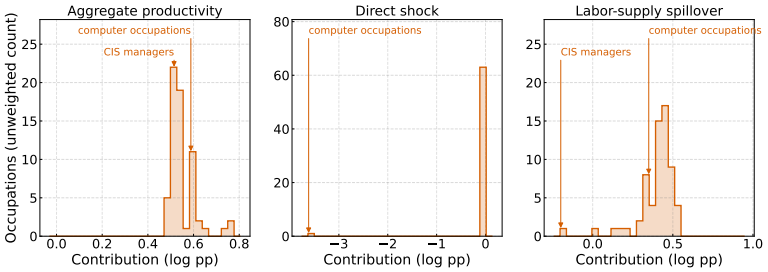
Figure A.24 summarizes the differences in the cross-occupation incidence of the shock across the two models. While the two models have similar predictions for skill prices in most occupations, the static benchmark predicts a substantially larger decline in the skill price for computer occupations, in addition to the smaller increase in the skill price for CIS managers. Accounting for young workers' endogenous skill accumulation decisions thus leads to a larger reallocation of labor away from computer occupations, which cushions the fall in the skill price for computer occupations while simultaneously creating skill scarcity and encouraging

[65] Figure A.23 separates labor-supply spillovers into contributions from the common wage increase and from deviations from the initial-wage-bill-weighted average. The relative-wage contribution is largest for computer managers in the learning economy and most negative in the static benchmark. This difference is driven primarily by the computer-specific wage contribution.

Figure 14: First-order decomposition of occupational skill price responses to computer automation



(a) Baseline model



(b) Static benchmark

Notes: Histograms show the cross-occupation distributions of the aggregate-output, direct labor-demand, and cross-occupation labor-supply contributions to the first-order change in occupational skill prices. Contributions are expressed in log percentage points: one point corresponds to a contribution of 0.01 to the change in the log skill price. Figure A.23 separates labor-supply spillovers into contributions from the common wage increase and from deviations from the (initial) wage-bill-weighted average.

additional automation further up the career ladder.

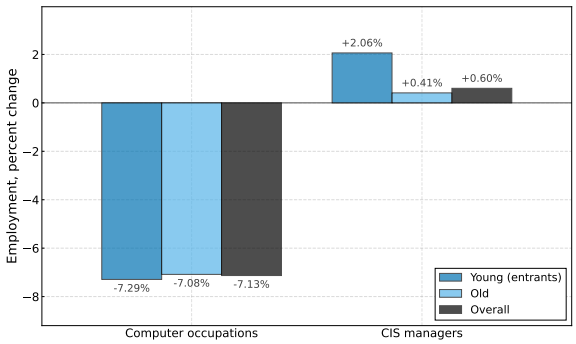
(Long-run) Worker-level incidence. Next, I use the model to examine the long-run incidence of the automation shock on different types of workers. I distinguish the effects of equilibrium wage adjustment from the gains that individual workers can obtain by adjusting their own occupational choices. I conduct this exercise in the learning model by comparing the expected lifetime utility for young workers entering the post-shock steady state. I first re-solve for the equilibrium wage response to the computer automation shock, holding labor supply fixed in every occupation and isolating the labor demand effects of the shock. I then evaluate expected welfare changes for each worker type at these wages while maintaining their pre-shock occupational choices, where welfare changes are expressed as the equivalent percentage change in consumption in every period of life.[66]

Figure 16 shows that all types of workers gain, on average, from the automation shock in the new steady state. However, this exercise also shows that long-run equilibrium skill price movements substantially cushion the shock for the engineering worker type, which is the group that is most exposed to the shock. Maintaining their original career choices, this group of workers experiences a consumption-equivalent *loss* of 0.63 percent under fixed-labor wages on average, but a gain of 0.23 percent at long-run equilibrium wages, while all other worker types enjoy smaller gains on average after allowing for labor supply adjustments.[67] Figure A.25 shows that

[66] I compute type-specific ex-ante consumption-equivalent welfare changes, computed from changes in expected lifetime utility for each worker type before the realization of idiosyncratic productivity shocks.

[67] For small shocks, the fact that virtually all gains come through skill price movements is a result of the envelope theorem. Because workers' initial career choices are optimal, the first-order change in expected lifetime utility for each group can be evaluated holding their original choices fixed.

Figure 15: Employment responses to the computer automation shock by age



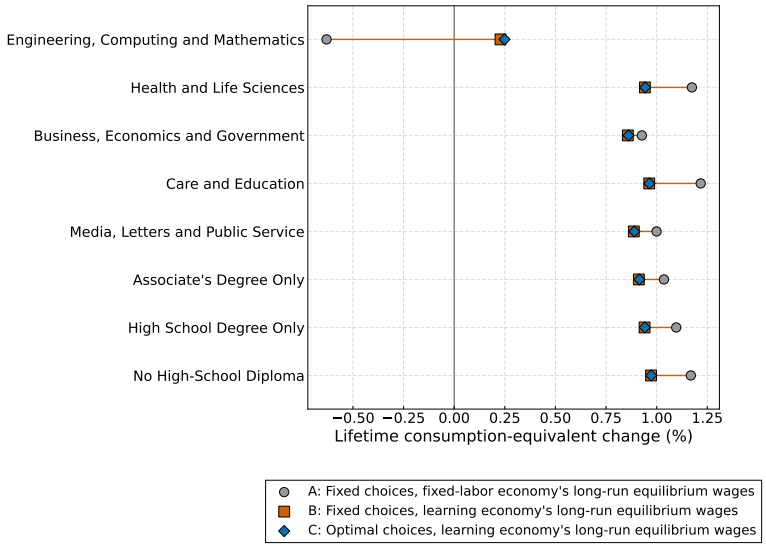
Notes: Figure reports employment responses in the baseline learning model following a reduction in the price of capital used in computer occupations. Young workers are in their initial career period. Older workers are in the subsequent post-learning model periods. Each age group's response is expressed as a percentage of its own initial employment in the occupation. The overall response thus weights the age-specific responses by their initial employment shares within the occupation.

long-run adjustments in labor supply reduce the decline in the skill price for computer occupations by almost three-quarters. The figure also shows that computer and information-systems managers are the only other occupation with a higher skill price in the learning economy's long-run equilibrium than in the fixed-labor reference. In all other occupations, increases in skill prices from expanding labor demand are dampened by expanding labor supply. Finally, Figure A.26 shows that the engineering worker type would also experience an expected welfare loss under the equilibrium wage changes computed in the static benchmark, as a model that assumes fixed skill portfolios absorbs less of the initial shock to computer occupations and misses much of the additional increase in the skill price for CIS managers.

Transitional dynamics. Finally, I use the model to examine how on-the-job learning shapes the economy's transitional dynamics following the initial automation shock. To perform this exercise, I consider an unanticipated and permanent reduction in the capital price for computer occupations. After the shock occurs, I assume that workers have perfect foresight over the future path of equilibrium aggregates. Figures 17 and 18 reveal that skill prices and capital shares, respectively, in computer occupations and CIS management follow non-monotonic paths. These responses reflect the interaction between forward-looking career choices and the gradual adjustment of skill supply as incumbent cohorts with predetermined skills are replaced.

Initially, incumbent workers reallocate from computer occupations toward CIS management, limiting the increase in managerial skill prices. Over time, however, reduced young entry into computer occupations means fewer workers acquire the skills useful for subsequent managerial work. The effects of this second force materialize only gradually, which explains why the increase in the skill price for CIS managers is delayed. The figure shows that two other occupations, computer support occupations and mathematical science occupations, also experience delayed increases and non-monotonicities in their skill price responses. In the quantified model, these two occupations are also dynamically linked to computer occupations, as they are a common entry-way into working in computer occupations, particularly for workers without a bachelor's degree. Finally, Figure 19 reveals that the two oldest incumbent cohorts of the engineering worker type that entered the labor market before the shock experience welfare *losses* on average, due to their relative immobility and the gradual equilibration of skill prices. These average losses, while modest, contrast with the long-run gains experienced by all worker types in the new steady state, as well as the large short-run

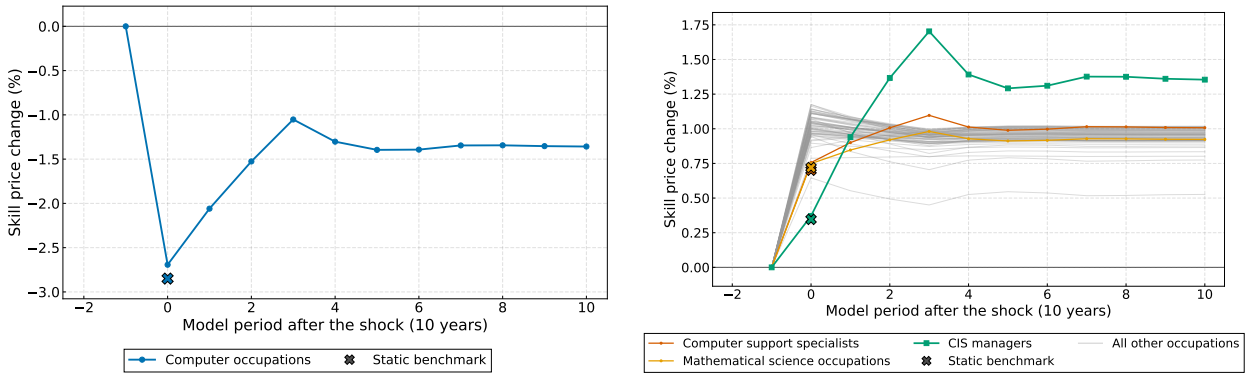
Figure 16: Incidence of the computer automation shock across worker types



Notes: The figure reports changes in young workers' expected lifetime utility by type, expressed as consumption-equivalent percentage changes. All cases evaluate the same automation shock in the baseline learning model. Point A evaluates workers' original career choices at the wages arising when effective labor supply is held fixed in each occupation. Point B evaluates those same choices at the learning economy's long-run equilibrium wages. Point C allows workers to adjust their career choices optimally at those equilibrium wages. The difference between A and B captures wage feedback from labor-supply adjustment; the difference between B and C captures the additional (average) welfare gain from adjusting career choices at those skill prices.

gains experienced by all other worker types along the transition.

Figure 17: Transitional effects of computer automation on skill prices

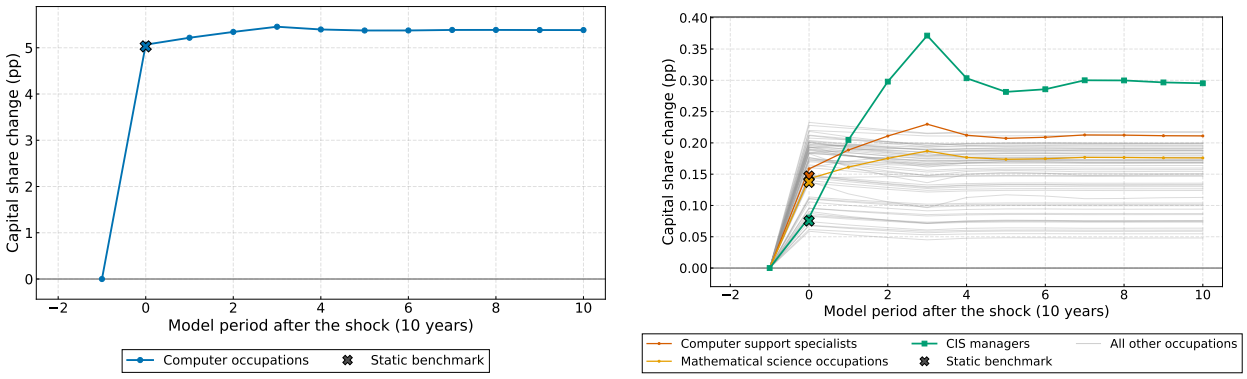


(a) Computer occupations

(b) CIS managers and other occupations

Notes: Figure shows skill-price responses to a permanent, unanticipated 22% reduction in the capital price for computer occupations at period zero. Solid lines show transition paths in the model with on-the-job learning. Panel (a) shows computer occupations. Panel (b) highlights computer and information systems (CIS) managers, computer support specialists, and mathematical science occupations; thin gray lines show the remaining occupations. Cross markers at period zero indicate the long-run static benchmark responses for computer occupations and CIS managers. Changes are measured in log percentage points relative to the initial steady state. Each model period represents ten years. Note the different vertical scales across panels.

Figure 18: Transitional effects of computer automation on capital shares

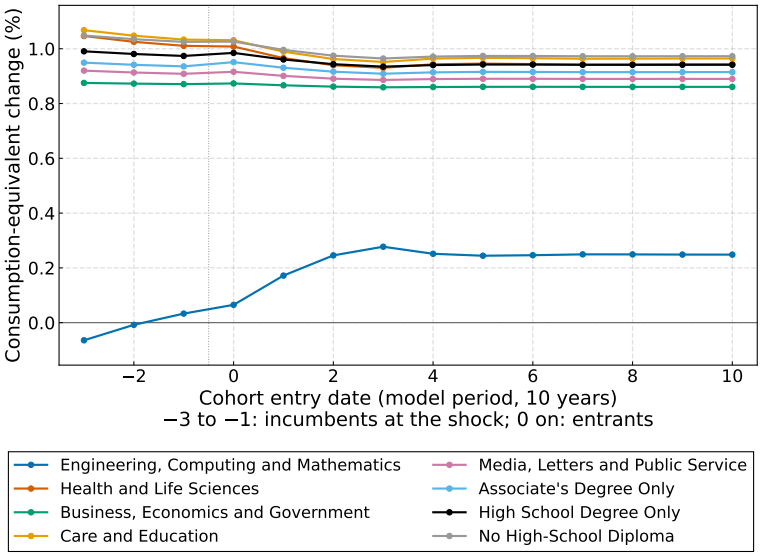


(a) Computer occupations

(b) CIS managers and other occupations

Notes: Figure shows capital-share responses to a permanent, unanticipated 22% reduction in the capital price for computer occupations at period zero. Solid lines show transition paths in the model with on-the-job learning. Panel (a) shows computer occupations. Panel (b) highlights computer and information systems (CIS) managers, computer support specialists, and mathematical science occupations; thin gray lines show the remaining occupations. Cross markers at period zero indicate the long-run static benchmark responses for computer occupations and CIS managers. Changes are measured in percentage points relative to the initial steady state. Each model period represents ten years. Note the different vertical scales across panels.

Figure 19: Welfare effects of computer automation by worker type and cohort



Notes: Figure reports consumption-equivalent welfare changes following a permanent, unanticipated 22% reduction in the capital price for computer occupations at period zero. The horizontal axis indicates the cohort's labor-market entry date, with each model period representing ten years. Negative dates identify workers already in the labor market when the shock occurs; their welfare changes are evaluated over their remaining lifetimes. For cohorts entering at period zero or later, changes are evaluated over their full lifetimes. All changes are relative to the initial steady state. For incumbents, expected utility changes are averaged over initial occupations within each type before computing consumption equivalents.

Alternative career paths. Appendix D.3 reports an analogue of Figure 13 for three additional examples of pairs of professional occupations that exhibit strong dynamic labor supply linkages in the quantified model, alongside uneven automation exposure to generative AI (legal support workers and lawyers, financial analysts and financial managers, accountants and financial managers).

6.3 Broad-based generative AI shock

In this section, I study the distributional consequences of a joint shock to the entire vector $\{A_j^K \cdot q_j\}$. To conduct this exercise, I map changes in capital efficiency to empirical measures of occupation-level automation exposure. However, an inherent challenge with this type of forward-looking exercise is that commonly used measures of automation exposure provide information on the feasibility of automation, rather than changes in the relative price of capital to labor in an occupation.

I therefore adopt a simple parameterization that sets $\tilde{A}_j^K \tilde{q}_j / A_j^K q_j = \exp(\kappa \cdot \text{Exposure}_j)$, where $\tilde{A}_j^K \tilde{q}_j$ denotes the resulting values of capital efficiency in the new steady state. Given my focus on distributional outcomes, I set $\kappa = 0.58$ to generate a 13% increase in (steady state) aggregate output in the baseline economy, using an illustrative benchmark motivated by economists' forecasts of AI's effects on GDP growth from Karger et al. (2026).[68] In Figure A.34, I show that the results are very similar under an alternative calibration that sets occupation-specific capital-price reductions to generate proportional unit-cost savings of $\text{Exposure}_j \cdot s_j^H \cdot 0.27$ for each occupation j, where unit costs are evaluated at skill prices in the initial steady state and 27% is motivated by the estimates of task-level cost savings from Acemoglu (2025).

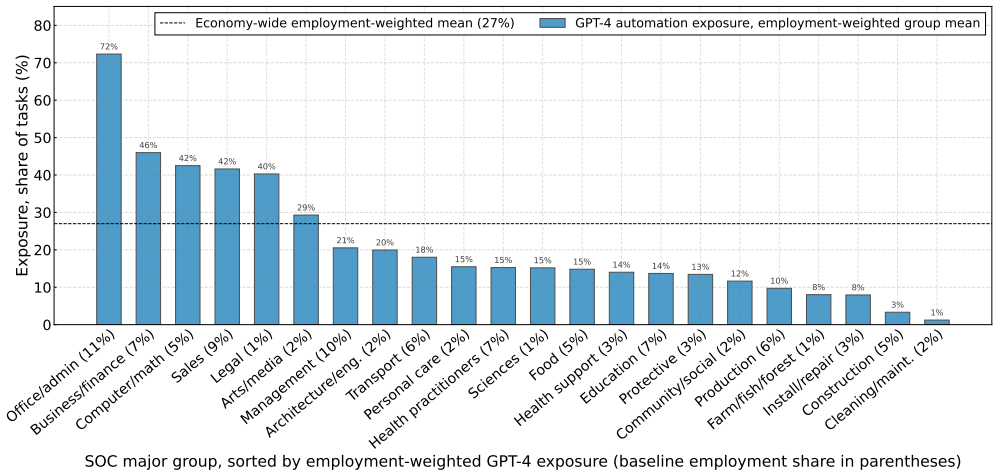
(Long-run) Occupation-level incidence. Figure 20 reports employment changes for broad occupational groups alongside each group's employment-weighted exposure to automation. The figure shows that employment contracts most sharply in highly exposed groups (office and administrative support, computer and mathematical, and business and financial occupations) and expands in managerial, service, and healthcare occupations. Relative to the static benchmark, the model with on-the-job learning generally predicts more pronounced employment reallocation away from exposed occupations toward less exposed occupations, reflecting how accounting for endogenous skill accumulation decisions increases the responsiveness of occupation-level labor supply. However, management occupations are a notable exception. Managerial employment expands less in the learning model than in the static benchmark, consistent with reduced entry into most professional occupations, which are among the occupations most exposed to AI-driven automation, limiting the supply of workers with skill portfolios suitable for managerial work.

Figure 21 plots long-run changes in occupational skill prices against exposure in the learning model and the static benchmark. The figure shows that in the model with on-the-job learning, skill prices are significantly less responsive to exposure. Accounting for forward-looking occupational choices and the resulting changes in skill supply thus provides a quantitatively important margin of long-run labor-supply adjustment. This dampening reduces the (initial) employment-weighted standard deviation of occupational skill-price responses by 43%, relative to the static benchmark.[69]

[68] In the baseline calibration of the model, all positive values of κ lead to aggregate output gains. Karger et al. (2026) compare annual GDP growth of 2.9% under moderate AI progress with a historical-growth benchmark of 2.4%. Maintaining these rates over 2025–2050 implies an output *level* difference of $(1.029/1.024)^{25} - 1 \approx 13\%$ in 2050.

[69] The baseline calibration generates relatively modest movements in relative skill prices. Figure A.33 shows how the relationship between occupational exposure and skill-price changes varies with the substitution elasticities η and λ, as well as the Fréchet parameter ν, which governs the responsiveness of labor supply. Intuitively, larger values of η and ν and smaller values of λ deliver larger movements in relative prices.

Figure 20: Automation exposure and employment responses by broad occupation group



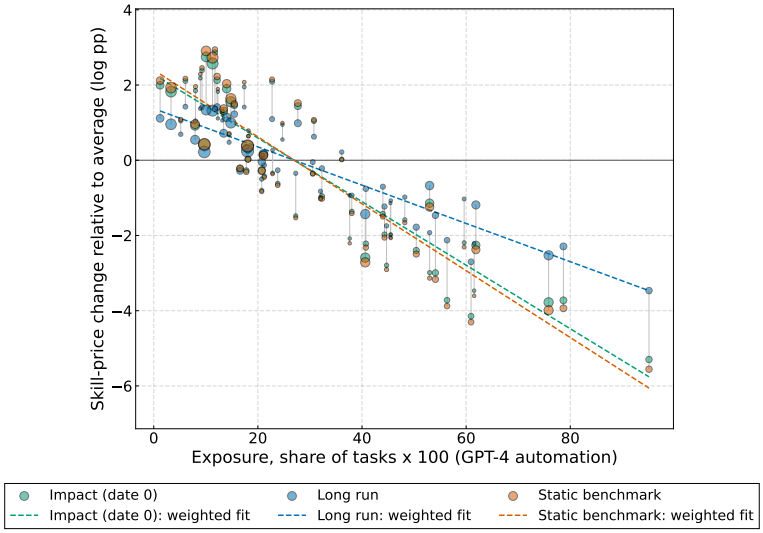
(a) Exposure to AI automation



(b) Long-run employment responses

Notes: Panel (a) reports the initial-employment-weighted mean GPT-4 automation exposure within broad occupational groups, expressed as the share of tasks exposed multiplied by 100. The dashed line indicates the economy-wide employment-weighted mean. Panel (b) reports percentage changes in each group's total employment following the broad-based automation shock, comparing the model with on-the-job learning and the static benchmark. Groups are ordered by decreasing exposure in both panels. Percentages in parentheses indicate each group's share of initial employment.

Figure 21: Occupational skill-price responses to the broad-based automation shock



Notes: Figure compares the impact (date zero) and long-run skill-price responses in the learning economy with the long-run response in the static benchmark, following the same permanent, unanticipated automation shock. Each point represents an occupation, with bubble sizes reflecting initial employment. The horizontal axis reports the percentage of tasks rated as automatable. The vertical axis reports 100 times the log skill-price change from the initial steady state, minus the initial-employment-weighted average change for the corresponding model and date. Thus, negative values indicate below-average skill-price growth, not necessarily an absolute decline. Thin gray lines connect the responses for the same occupation across the three series. Dashed lines are separate linear fits weighted by initial occupational employment.

(Long-run) Worker-level incidence. Figure 22 reports changes in young workers' expected lifetime utility in the new steady state, separately for each worker type. The figure shows that all types of workers gain from the broad-based AI-driven automation shock, with average welfare gains ranging from approximately 5.9% to 7.4%. The figure also shows that long-run adjustments in skill prices increase average welfare gains for the business and engineering type workers, while reducing the gains accruing to the health, care, and education type workers, as well as workers without a high-school diploma. These gradual adjustments in skill prices narrow the gap between the largest and smallest type-specific gains from approximately 4.8 to 1.6 percentage points and reflect the gradual reallocation of workers away from professional occupations toward healthcare and service occupations.[70]

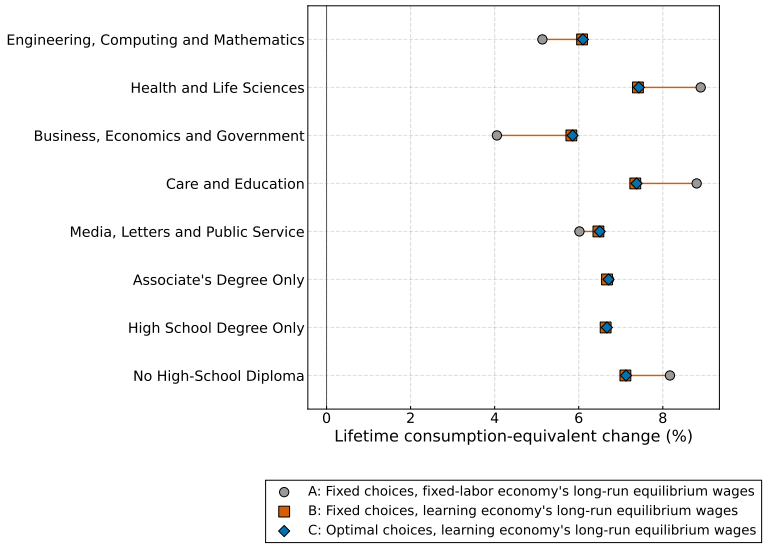
Figure A.31 plots average percentage changes in lifetime income across types, where types are ordered by their average lifetime income in the initial steady state. Overall, I find that the effects of the AI-driven automation shock on lifetime income inequality are relatively modest, similar to recent work of Freund and Mann (2026) that focuses on the impacts of generative AI-driven automation on (static) wage inequality.[71]

Transitional dynamics. Finally, I study the dynamic response of this economy to the automation shock. Figures A.32a and A.32b show the dynamic adjustment of occupational skill prices and the resulting welfare impacts on different cohorts of workers, respectively. On impact, changes in skill prices in the learning model are very close to those of the static benchmark, reflecting the fixed skill portfolios of incumbent workers. Quantitatively, the slope relating occupational log skill-price changes to automation exposure is

[70] Figure A.30 shows that larger values of κ increase welfare gains for all types while widening the differences in gains across types. The figure also plots welfare gains under the equilibrium wage changes computed in the static benchmark model.

[71] Smeets et al. (2025) model AI as a direct shock to the returns to skills in different fields and find a much larger, progressive effect on lifetime income inequality.

Figure 22: Incidence of broad-based AI shock across worker types



Notes: The figure reports changes in young workers' expected lifetime utility by type, expressed as consumption-equivalent percentage changes. All cases evaluate the same broad-based automation shock in the baseline learning model. Point A evaluates workers' original career choices at the wages arising when effective labor supply is held fixed in each occupation. Point B evaluates those same choices at the learning economy's long-run equilibrium wages. Point C allows workers to adjust their career choices optimally at those equilibrium wages. The difference between A and B captures wage feedback from labor-supply adjustment; the difference between B and C captures the additional private gain from career adjustment at those wages. Holding original choices fixed preserves baseline occupational switches and sorting on productivity draws rather than prohibiting mobility altogether.

only about 5% smaller in magnitude on impact than in the static benchmark. Its magnitude then falls by approximately 40% between impact and the new steady state, highlighting how adjustments in career choices and the workforce's skill composition gradually dampen the relative skill-price effects of automation.

The compression of skill prices and the welfare gains across types in the learning model emerges slowly, taking place through gradual changes in the workforce's skill composition. However, unlike the previous quantitative exercise, I find that all worker types experience welfare gains throughout the transition, given the large aggregate productivity gains accompanying the automation shock. Still, the distribution of gains varies meaningfully across cohorts. Later entrants of the engineering and business types gain more than incumbent cohorts, whereas incumbents of the health science and care types gain more than entrants in the new steady state. These patterns reflect how successive cohorts of workers adjust their occupational choices to acquire skills better suited to the post-automation pattern of labor demand, ultimately spreading the long-run gains more evenly across workers with different initial skill sets.

Limitations. The quantitative results should be interpreted in light of the limitations discussed below, several of which I plan to address in future work.

First, substantial uncertainty surrounds the future capabilities and costs of generative AI technologies. When studying the transitional dynamics following an automation shock, I abstract from workers' uncertainty about future technological change and the value of different career paths. Such uncertainty could affect their occupational choices and the welfare consequences of technological change.

Second, the counterfactuals hold labor-supply primitives fixed, including initial skills, amenities, and the

on-the-job learning technology. Although skill accumulation responds endogenously to occupational choices, learning opportunities within each occupation remain unchanged by the shock. These exercises therefore do not capture how AI-induced changes in job content might alter the skills workers acquire on the job. For example, automating particular tasks within occupations could eliminate opportunities to develop certain skills. Similarly, I do not study how AI might change the skill requirements of an occupation, which is the focus of several recent studies (Althoff and Reichardt (2026), Freund and Mann (2026), Hosseini Maasoum and Lichtinger (2026a)). Moreover, AI-driven job transformation could also directly change the amenity value of working in a given occupation, another margin held fixed in the analysis.

Third, the model abstracts from explicit occupational entry and switching costs, licensing constraints, and other mobility frictions, likely leading it to overstate workers' abilities to reallocate across occupations. Conversely, reflecting the focus of this paper on initial career choices, the baseline calibration restricts skill accumulation to the first decade of a worker's career. This assumption is likely to understate the ability of incumbent workers to adapt to technological change by acquiring new skills later on in their careers.

Finally, the parsimonious production structure restricts the types of complementarities that can arise across different types of workers, occupations, and tasks. Richer production relationships, such as complementarities between junior and senior workers within occupations or industries, could further amplify or offset the labor-supply complementarities that are the focus of this paper.

7 Conclusion

This paper develops a dynamic general equilibrium model of occupational choice with endogenous human capital accumulation. Combining the structure of the occupational choice model with large-scale résumé data, I recover a flexible, reduced-form on-the-job learning technology that rationalizes early-career mobility across granular occupations. I show that the recovered learning estimates are closely associated with, but not fully explained by, observable measures of knowledge proximity and industry overlap, consistent with the accumulation of domain-specific human capital. The quantified model rationalizes the occupational career ladders that I document in the data and shows how such ladders can generate longer-run complementarities in labor supply by which higher wages in one occupation increase the labor supplied to another.

I use the quantitative model to study the potential labor market impacts of generative AI-driven automation shocks within and across occupational career ladders. The quantitative experiments show how complementarities in labor supply can propagate an automation shock further up the occupational career ladder. For a more broad-based automation shock, however, I find that accounting for endogenous skill accumulation choices primarily dampens the dispersion in occupational skill prices and evens the distribution of gains across workers with different types of skills. For both automation shocks, I find that the full effects of the automation shock unfold gradually as successive cohorts adjust their initial career choices, so that the incidence on incumbent workers can differ substantially from that of labor market entrants in the new steady state, particularly for the workers whose initial skill profiles are most exposed to the initial automation shock.

My analysis suggests multiple avenues for future research. First, financial frictions that limit workers' ability to accept lower current earnings in exchange for better learning opportunities could dampen the aggregate gains from automation directed at training occupations. Recent empirical work has demonstrated how financial constraints can distort the educational and early-career occupational choices of young individuals

(Hampole (2024), Luo and Mongey (2019)). Future work could build on these findings to examine how financial frictions shapes occupational choices in response to automation and its aggregate consequences. Another promising direction is to incorporate endogenous educational choices, providing young workers with an additional margin of adjustment in response to changing returns across career paths. Recent evidence from Denmark suggests that field-of-study choices are an empirically relevant margin of labor-supply adjustment (Smeets et al. (2025)). On the labor-demand side, endogenizing the development of automation technologies would capture how workers' skill investments feed back into the incentives for technological innovation. In the current model, firms endogenously adjust their use of automation, but improvements in the underlying technologies are exogenous. Finally, historical episodes of automation offer an opportunity to test the model's empirical prediction that automation shocks to early-career occupations or tasks encourages future automation in occupations further up the career path.

References

The numbers at the end of every reference link to the pages citing the reference.

Acemoglu, Daron, "The simple macroeconomics of AI," *Economic Policy*, 01 2025, *40* (121), 13–58. 44

_ and David Autor, "Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 2011, *4*, 1043–1171. 4, 7, 53

_ and Pascual Restrepo, "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, May 2019, *33* (2), 3–30. 16

_ and _ , "Tasks, Automation, and the Rise in U.S. Wage Inequality," *Econometrica*, 2022, *90* (5), 1973–2016. 12, 15, 16, 32, 62

Afrouzi, Hassan, Andres Blanco, Andrés Drenik, and Erik Hurst, "Automation, Learning, and Career Dynamics," Working Paper 35157, National Bureau of Economic Research April 2026. 1, 5

Althoff, Lukas and Hugo Reichardt, "Task-Specific Technical Change and Comparative Advantage," NBER Working Paper 35353, National Bureau of Economic Research June 2026. 2, 3, 5, 16, 48

Asriyan, Vladimir, Hugo Reichardt, and Sandro Shelegia, "Immiserizing Automation," 2026. Working paper. 5

Atalay, Enghin, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum, "The Evolution of Work in the United States," *American Economic Journal: Applied Economics*, April 2020, *12* (2), 1–34. 4

Autor, David H., Frank Levy, and Richard J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 11 2003, *118* (4), 1279–1333. 4

_ , Lawrence F. Katz, and Melissa S. Kearney, "The Polarization of the U.S. Labor Market," *American Economic Review*, May 2006, *96* (2), 189–194. 4

Ben-Porath, Yoram, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 1967, *75* (4, Part 1), 352–365. 16

Brynjolfsson, Erik, Bharat Chandar, and Ruyu Chen, "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence," Technical Report, Stanford Digital Economy Lab August 2026. 5

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel, "Changes in Between-Group Inequality: Computers, Occupations, and International Trade," *American Economic Journal: Macroeconomics*, April 2019, *11* (2), 348–400. 26, 32, 36

Caselli, Francesco and Alan Manning, "Robot Arithmetic: New Technology and Wages," *American Economic Review: Insights*, June 2019, *1* (1), 1–12. 18

Caunedo, Julieta, David Jaume, and Elisa Keller, "Occupational Exposure to Capital-Embodied Technical Change," *American Economic Review*, June 2023, *113* (6), 1642–85. 26, 32, 36

Chinoy, Sahil and Martin Koenen, "Political Sorting in the U.S. Labor Market: Evidence and Explanations," December 2024. Job Market Paper, Harvard University. 53

Cortes, Guido Matias and Giovanni Gallipoli, "The Costs of Occupational Mobility: An Aggregate Analysis," *Journal of the European Economic Association*, 04 2018, *16* (2), 275–315. 26

Deming, David, "Why do Wages Grow Faster for Educated Workers?," *Journal of Labor Economics*, 2025, *0* (ja). 2, 7

Dingel, Jonathan I. and Felix Tintelnot, "Spatial Economics for Granular Settings," *Econometrica*, 2026, *94* (2), 407–464. 29

Dix-Carneiro, Rafael, "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 2014, *82* (3), 825–885. 4

Dorn, David, Florian Schoner, Moritz Seebacher, Lisa Simon, and Ludger Woessmann, "Multidimensional Skills on LinkedIn Profiles: Measuring Human Capital and the Gender Skill Gap," March 2026. Working paper, March 28, 2026. 6, 7, 53

Dvorkin, Maximiliano A. and Alexander Monge-Naranjo, "Occupation Mobility, Human Capital and the Aggregate Consequences of Task-Biased Innovations," Working Paper 2019-013, Federal Reserve Bank of St. Louis April 2019. 2, 4, 13, 36, 60

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock, "GPTs are GPTs: Labor market impact potential of LLMs," *Science*, 2024, *384* (6702), 1306–1308. 8, 9, 11, 56

Fan, Tianyu, "The Labor Market Incidence of New Technologies," 2025. 8, 16, 33, 63

Freund, Lukas B. and Lukas F. Mann, "Job Transformation, Specialization, and the Labor Market Effects of AI," March 2026. Working paper, version dated March 18, 2026. 8, 16, 36, 46, 48, 83

Friebel, Guido, Yao Huang, Jin Li, Soumitra Shukla, and Yunchou Zhang, "Pyramids, Diamonds, and Oscillations: AI and the Structure of Internal Labor Markets," Working Paper 6570519, SSRN April 2026. 5

Garicano, Luis and Luis Rayo, "Training in the Age of AI: A Theory of Career Viability," Working Paper May 2026. May 14, 2026. 5

Gathmann, Christina and Uta Schönberg, "How General Is Human Capital? A Task-Based Approach," *Journal of Labor Economics*, 2010, *28* (1), 1–49. 4

Goos, Maarten, Alan Manning, and Anna Salomons, "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, August 2014, *104* (8), 2509–26. 4, 36

Grigsby, John, "Does Skill Heterogeneity Affect Aggregate Employment-Wage Comovements?," 2025. 33

Hampole, Menaka, Dimitris Papanikolaou, Lawrence D.W Schmidt, and Bryan Seegmiller, "Artificial Intelligence and the Labor Market," Working Paper 33509, National Bureau of Economic Research February 2025. 6, 7

Hampole, Menaka V., "Financial Frictions and Human Capital Investments," March 2024. Working Paper. 49

Hawkins, William, Ryan Michaels, and Jiyoon Oh, "The Joint Dynamics of Capital and Labor at the Plant Level," *Working paper*, 2015. 36

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow, "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 2019, *87* (5), 1439–1474. 26

Humlum, Anders, "Robot Adoption and Labor Market Dynamics," 2021. Unpublished manuscript. 4

Ide, Enrique, "Automation, AI, and the Intergenerational Transmission of Knowledge," 2026. 1, 5

Karger, Ezra, Otto Kuusela, Jason Abaluck, Kevin A Bryan, Basil Halperin, Todd R Jones, Connacher Murphy, Philip Trammell, Matt Reynolds, Dan Mayland, Ria Viswanathan, Ananaya Mittal, Rebecca Ceppas de Castro, Josh Rosenberg, and Philip Tetlock, "Forecasting the Economic Effects of AI," Working Paper 35046, National Bureau of Economic Research April 2026. 44

Lise, Jeremy and Fabien Postel-Vinay, "Multidimensional Skills, Sorting, and Human Capital Accumulation," *American Economic Review*, August 2020, *110* (8), 2328–76. 4

Luo, Mi and Simon Mongey, "Assets and Job Choice: Student Debt, Wages and Amenities," Working Paper 25801, National Bureau of Economic Research May 2019. 49

Maasoum, Seyed Mahdi Hosseini and Guy Lichtinger, "Generative AI and Occupational Entry Barriers: The Labor-Supply Channel of Technological Change," SSRN Working Paper 6059674, SSRN 2026. 9, 11, 48, 56, 59

_ and _ , "Generative AI as Seniority-Biased Technological Change: Evidence from U.S. Résumé and Job Posting Data," SSRN Working Paper 5425555, SSRN 2026. 5, 6, 7

Smeets, Valerie, Lin Tian, and Sharon Traiberman, "Field Choice, Skill Specificity, and Labor Market Disruptions," Working Paper 33728, National Bureau of Economic Research April 2025. 23, 27, 46, 49

Traiberman, Sharon, "Occupations and Import Competition: Evidence from Denmark," *American Economic Review*, December 2019, *109* (12), 4260–4301. 3, 4, 23

Appendix

A Data appendix

A.1 Occupational task content measures from Acemoglu and Autor (2011)

Table A.1: Acemoglu–Autor occupational task content measures

Task content measure	O*NET code	O*NET component
Routine cognitive	4.C.3.b.7	Importance of repeating the same tasks
	4.C.3.b.4	Importance of being exact or accurate
	4.C.3.b.8	Structured v. Unstructured work (reverse)
Routine manual	4.C.3.d.3	Pace determined by speed of equipment
	4.A.3.a.3	Controlling machines and processes
	4.C.2.d.1.i	Spend time making repetitive motions
Non-routine cognitive: analytic	4.A.2.a.4	Analyzing data/information
	4.A.2.b.2	Thinking creatively
	4.A.4.a.1	Interpreting information for others
Non-routine cognitive: interpersonal	4.A.4.a.4	Establishing and maintaining personal relationships
	4.A.4.b.4	Guiding, directing and motivating subordinates
	4.A.4.b.5	Coaching/developing others
Non-routine manual	4.A.3.a.4	Operating vehicles, mechanized devices, or equipment
	1.A.2.a.2	Manual dexterity
	1.A.1.f.1	Spatial orientation
	4.C.2.d.1.g	Spend time using hands to handle, control or feel objects, tools or controls

Notes: Table lists the O*NET components underlying the five task content measures of Acemoglu and Autor (2011). I construct the measures using O*NET version 27.2. "Reverse" indicates that the component is reverse-coded.

A.2 Revelio data construction

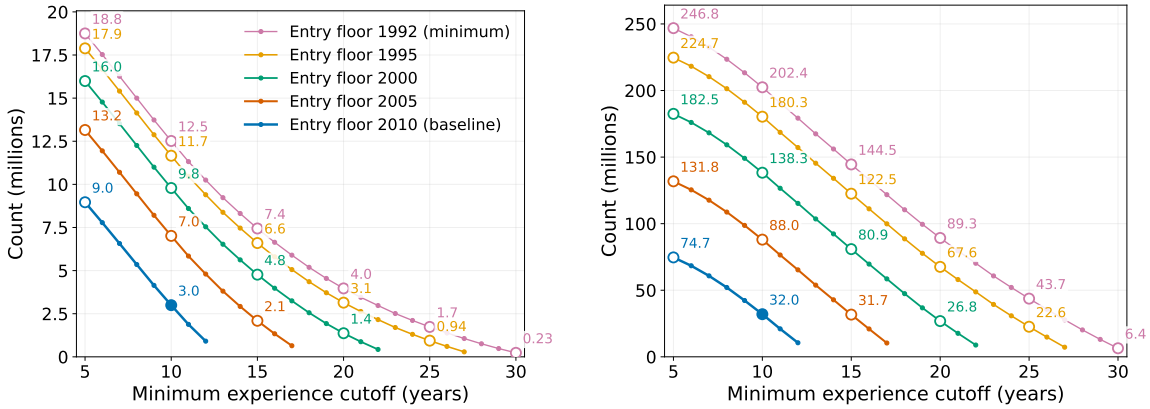
Basic data cleaning. I drop the small number of positions with a reported end date that lies before the reported start date. For positions with no listed end date, I treat the position as ongoing as of the data refresh date on WRDS, but all results are robust to imputing an end date as the start date of the next position held as in Chinoy and Koenen (2024) and Dorn et al. (2026). Following Dorn et al. (2026), I exclude individuals (i) whose first reported position begins more than five years after their latest reported degree completion date or (ii) whose first position begins before (inferred) age fourteen. I exclude individuals whose first reported position begins after (inferred) age thirty-one, given my focus on early career workers. I exclude users who ever report holding more than five ongoing positions at a time, as it is not possible to reliably assign a main job for these users.

Defining a main job. To define the main job, I consider all jobs held longer than 90 days in the calendar year, after excluding any positions with titles or descriptions that suggest non-employment status (e.g.,

"retired professor" or "law student"). I then define the job with the highest seniority index assigned by Revelio as the main job. Revelio's seniority index is an ordinal measure ranging from 1 (entry-level or intern) to 7 (senior executive), inferred from the job title, employer and industry characteristics, and the worker's career history. Ties are resolved by selecting the job held for the longest duration in that year, then the later start date and, finally, the lowest Revelio position identifier. Results are robust to alternative definitions of the main job.

Mapping job titles to 2018 SOC occupations. First, I correct systematic misclassifications at the role level, such as attorney roles assigned to non-lawyer occupations. Second, I refine individual position assignments using standardized matches to the BLS 2018 SOC Direct Match Title File and targeted rules for ambiguous legal and clinical roles (such as "partner" and "associate"). Positions not affected by these corrections retain their Revelio role assignment. For some licensed occupations, I additionally exclude assignments unsupported by the worker's reported qualifications or reported job title. Additionally, I drop the small number of positions that are missing both a job title and job description. Revelio must infer occupations for these individuals from other information on their profile, sometimes using descriptions that were provided in the employer description. In practice, I find that these positions are often mapped to generic, catch-all roles like "Office Worker".

Figure A.1: Revelio sample size



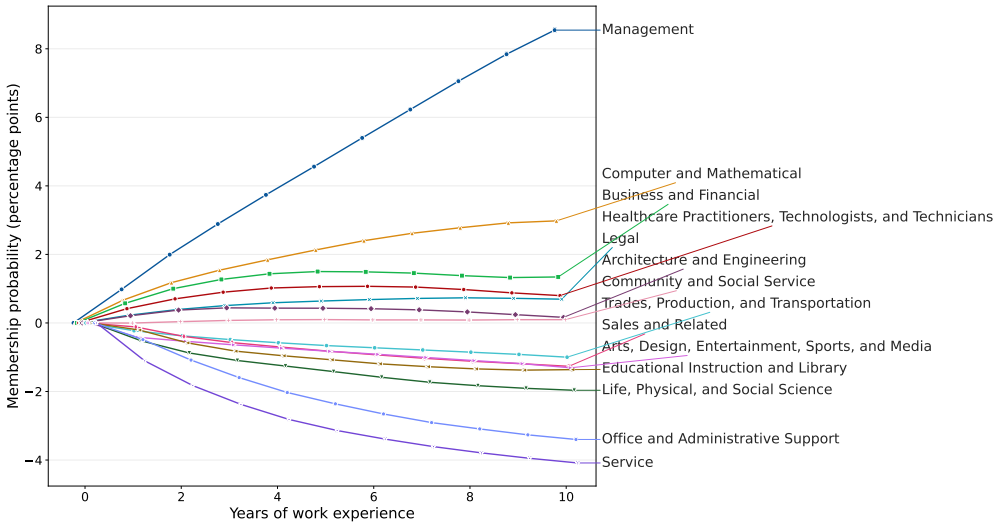
Panel A: Workers Panel B: Observations

Notes: Panel A shows the number of unique users in the dataset as I vary the entry floor and minimum experience cutoffs. Panel B shows the resulting number of observations (workers × years) associated with these users.

A.3 Additional figures and tables for Section 2

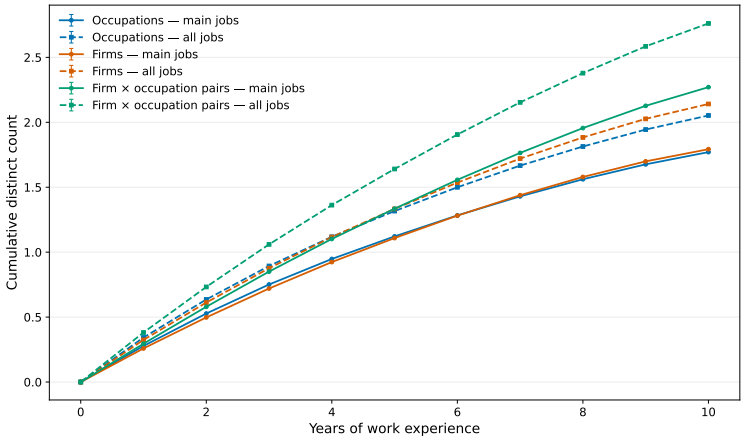
A.3.1 Patterns of occupational mobility

Figure A.2: Patterns of early-career occupational mobility



Notes: Figure plots within-worker changes in the probability of employment in each broad occupational group over the first ten years of potential experience. Each series reports coefficients from a separate regression of an occupational-group membership indicator on annual experience indicators and worker fixed effects. Standard errors are clustered at the user level, and 95% confidence intervals are shown.

Figure A.3: Early-career occupational and firm mobility



Notes: Figure plots within-worker changes in the cumulative number of distinct occupations, firms, and firm–occupation pairs observed over the first ten years of potential experience. Solid lines use main-job histories; dashed lines use all-job histories. Each series reports coefficients from a separate regression on annual experience indicators and worker fixed effects, relative to zero years of experience. Occupations are defined using detailed 2018 SOC codes. Results are similar for the level of occupational aggregation used to construct the empirical transition networks. Standard errors are clustered at the user level, and 95% confidence intervals are shown.

A.3.2 Details on exposure scores

Task ratings and aggregation. I construct the GPT-4 measures from the task-level automation ratings of Eloundou et al. (2024), and the GPT-5 measures from the updated ratings of Hosseini Maasoum and Lichtinger (2026a). Both sources rate tasks from O*NET version 27.2. The automation rubric assigns tasks to five ordered categories, T0 through T4, ranging from no automation exposure to low, moderate, high, and full automation exposure. The GPT-5 series uses the updated GPT-5.2 ratings. These measures describe the technological feasibility of automating tasks, rather than observed adoption or realized employment displacement. Differences in the rating rubrics also mean that the comparison does not isolate changes in model capabilities.

For model m and occupation o, let T_{ot}^{m} denote the rating of task t, and let ω_{ot} be its normalized weight within the occupation. I give core tasks twice the weight of supplemental tasks and use the same original task weights for both models.[72] The two automation measures are

$$A_o^{m,\text{binary}} = \sum_{t \in \mathcal{T}_o} \omega_{ot} \, \mathbb{1}\{T_{ot}^{m} \in \{\text{T3}, \text{T4}\}\}, \tag{A.1}$$

$$A_o^{m,\text{continuous}} = \sum_{t \in \mathcal{T}_o} \omega_{ot} \, s(T_{ot}^{m}), \tag{A.2}$$

where $\sum_t \omega_{ot} = 1$ and $(s(\text{T0}), s(\text{T1}), s(\text{T2}), s(\text{T3}), s(\text{T4})) = (0, 0.25, 0.5, 0.75, 1)$. The *binary* measure is therefore the weighted share of tasks rated as highly or fully automatable; "binary" refers to the task-level indicator, not the resulting occupation-level share. The *continuous* measure retains distinctions among all five categories by averaging their numerical scores. It is an index of automation exposure, not a literal estimate of the fraction of work that can be automated.[73]

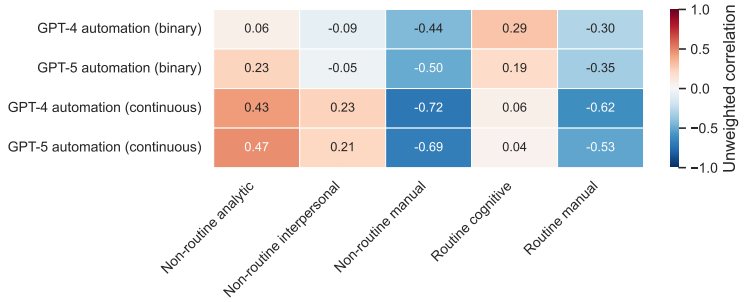
I map the O*NET-SOC scores to detailed 2018 SOC occupations. When an exact occupation-level O*NET code ending in `.00` has a non-missing score, I use that score; otherwise, I average the available scores of its specialty occupations. This rule is applied separately to each measure, and missing scores are not treated as zero. The distribution and wage figures below use exposure in its original units on $[0, 1]$, rather than the standardized units used in the experience-profile regressions.

Relationship to occupational task content. Figure A.4 relates the exposure measures to the five Acemoglu–Autor task content measures described in Table A.1. Each cell is a Pearson correlation across detailed 2018 SOC occupations, with equal weight on each occupation and using the available observations for that particular pair of measures. Comparing the binary and continuous automation rows shows how the relationship with task content depends on whether only high/full automation ratings count or intermediate ratings also receive positive scores.

[72] Tasks with missing original task-type classifications receive the supplemental-task weight. Holding these weights fixed avoids changes due to the updated source's reclassification of some tasks as core.

[73] For the small number of updated GPT-5 task ratings containing multiple categories, I average the listed category scores for the continuous measure and retain the source's supplied binary coding, which assigns these cases zero.

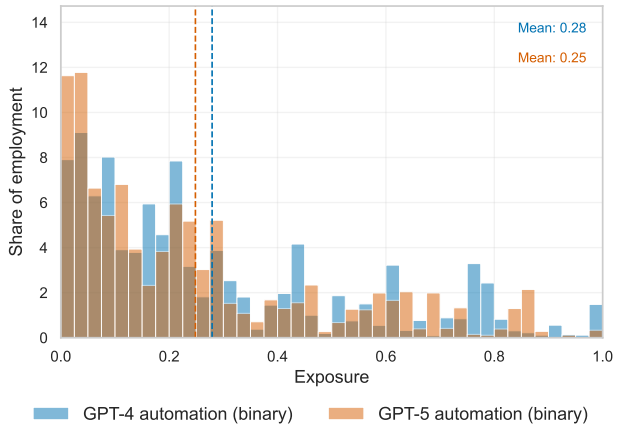
Figure A.4: Exposure measures and occupational task content



Notes: Entries are unweighted Pearson correlations across detailed 2018 SOC occupations between the exposure measure in each row and the Acemoglu–Autor task content measure in each column. Each correlation uses occupations with non-missing values for both measures, so samples can differ across cells. The task content measures use O*NET version 27.2. Binary and continuous automation exposure are defined in Equations (A.1) and (A.2).

Distribution across occupations. Figure A.5 compares the GPT-4 and GPT-5 binary automation measures across detailed occupations. I restrict to occupations with both scores and positive, observed 2022 OEWS employment. The histograms use common bins on $[0, 1]$ and weight each occupation by its share of employment in this sample; each bar therefore reports the percentage of covered employment in the corresponding exposure interval. Dashed lines mark the employment-weighted mean under each rating system.

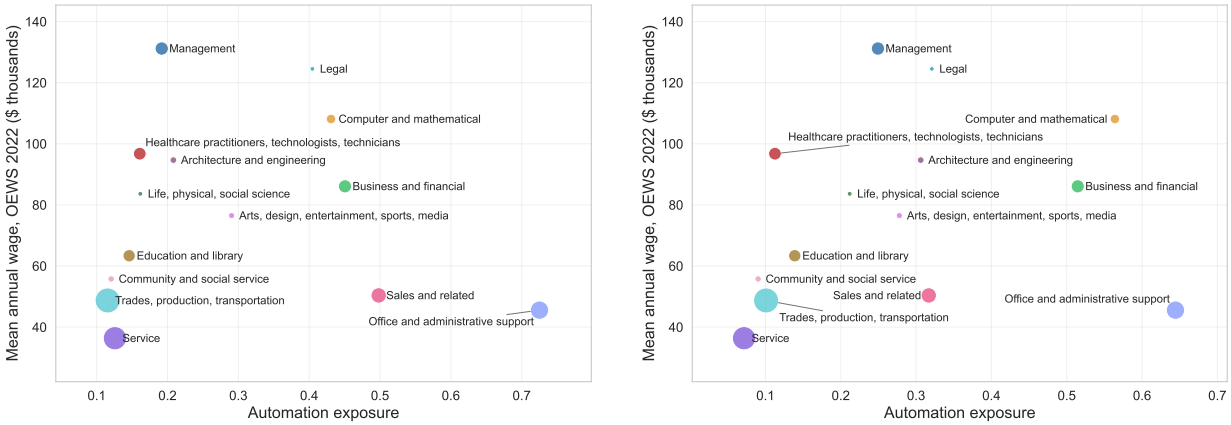
Figure A.5: Employment-weighted distribution of automation exposure



Notes: Figure compares binary GPT-4 and GPT-5 automation exposure, measured as the core/supplemental-weighted share of tasks rated T3 or T4. Bars report percentages of 2022 OEWS employment among detailed 2018 SOC occupations with both exposure scores and positive employment weights. Both histograms use the same bins. Dashed lines show the corresponding employment-weighted means.

Exposure and wages. Figure A.6 aggregates detailed occupations to broad occupational classes. Within each class, I average exposure and mean annual wages using detailed occupations' 2022 OEWS employment as weights, retaining non-missing observations for each average. Each bubble's area is proportional to the class's share of total employment across the included classes. The two panels use the same wage data and employment weights and differ in whether exposure is based on GPT-4 or GPT-5 ratings. These are descriptive comparisons across occupational classes, not estimates of the effect of automation on wages.

Figure A.6: Automation exposure and wages across occupational classes



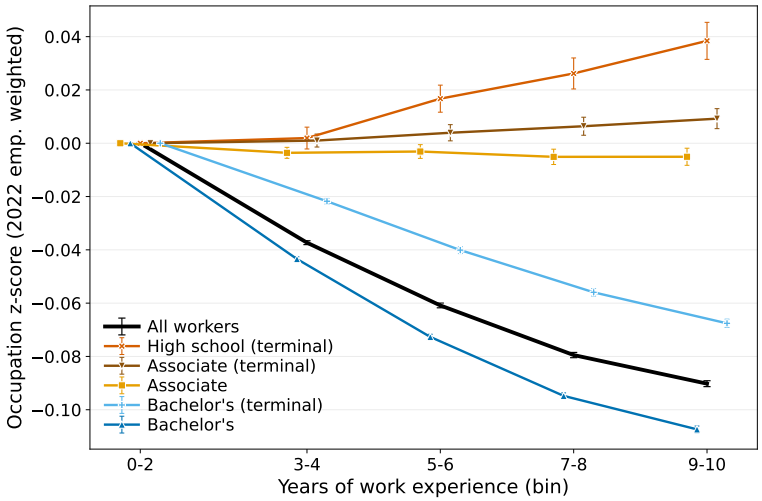
(a) GPT-4 binary automation exposure

(b) GPT-5 binary automation exposure

Notes: Each point is a broad occupational class. The horizontal axis reports mean binary automation exposure and the vertical axis reports mean annual wages in thousands of dollars, both averaged over detailed occupations using 2022 OEWS employment weights. Bubble area is proportional to the class's share of 2022 employment. Exposure and wages are averaged over the available observations for each measure. Panel (a) uses GPT-4 ratings and panel (b) uses GPT-5 ratings.

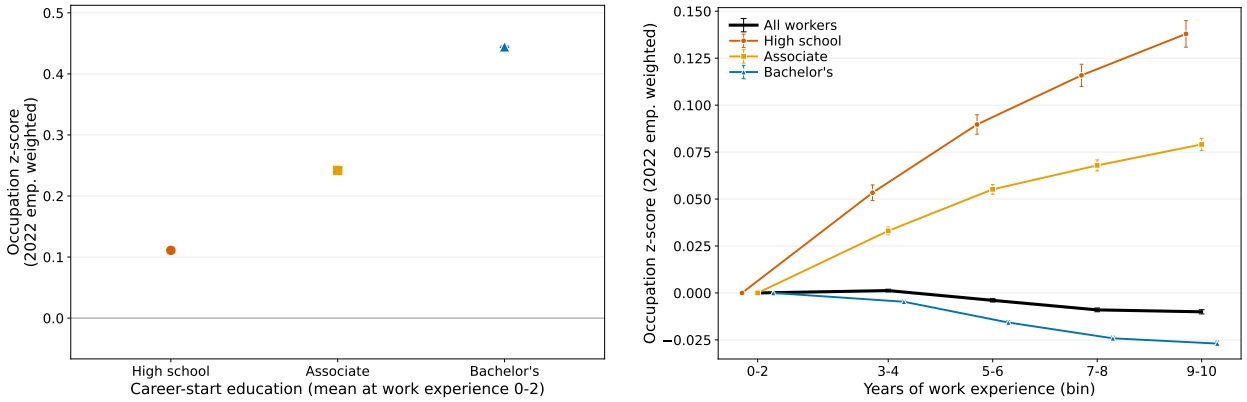
A.3.3 Additional automation exposure profiles

Figure A.7: Automation exposure by career-start education and terminal degree



Notes: Figure reports changes in standardized GPT-4 automation exposure with work experience by career-start education, together with the terminal-degree subsets. Terminal-degree users are those whose highest degree record matches the degree used to determine their labor market entry date; the high-school-start group is terminal by construction. Profiles are estimated using Regression (2), relative to 0–2 years of work experience. Standard errors are clustered at the user level, and 95% confidence intervals are shown.

Figure A.8: Automation exposure to generative AI: GPT-5 scores



(a) Average exposure at 0–2 years of experience

(b) Changes in exposure

Notes: Figure reproduces Figure 2 using the GPT-5 automation scores constructed by Hosseini Maasoum and Lichtinger (2026a). Panel (a) reports average exposure for the three career-start educational groups at 0–2 years of work experience. Panel (b) plots the estimated experience coefficients from Regression (2) using standardized GPT-5 exposure. The associate's- and bachelor's-start groups include users who subsequently complete higher degrees. Standard errors are clustered at the user level, and 95% confidence intervals are shown. Differences in the task-rating rubrics mean that comparisons with the GPT-4 results do not isolate changes in model capabilities.

B Model appendix

B.1 Task-based microfoundation for the CES production function in (3)

Following Dvorkin and Monge-Naranjo (2019), consider a unit continuum of tasks $x \in [0,1]$ within occupation j. Each task can be performed using effective labor or capital, which are perfect substitutes at the task level:

$$y_j(x) = A_j^H z_j^H(x) h_j(x) + A_j^K z_j^K(x) k_j(x), \qquad h_j = \int_0^1 h_j(x)\,dx, \quad k_j = \int_0^1 k_j(x)\,dx. \tag{A.3}$$

The task-level productivity draws z_j^H, z_j^K are independent unit-scale Fréchet variables with common shape $\eta - 1 > 0$: $\Pr(z_j^f \le z) = \exp(-z^{-(\eta-1)})$, $f \in \{H, K\}$. Their cross-sectional distribution across tasks equals this joint law. Thus, A_j^H and A_j^K are the scales of the two task productivity distributions, and task draws are fixed when these scales change. Within an occupation tasks are aggregated in a CES production function with elasticity of substitution σ satisfying $0 < \sigma < \eta$:

$$\begin{aligned} y_j &= \kappa_{\sigma,\eta} \cdot \left(\int_0^1 y_j(x)^{(\sigma-1)/\sigma}\,dx \right)^{\sigma/(\sigma-1)}, \\ \kappa_{\sigma,\eta} &= \Gamma\left(1 + \frac{1-\sigma}{\eta-1}\right)^{1/(1-\sigma)}. \end{aligned} \tag{A.4}$$

The constant $\kappa_{\sigma,\eta}$ is a normalization chosen to recover the main text's production function exactly. Note that the elasticity σ is distinct from λ, the elasticity across occupation outputs, and η, the resulting elasticity between capital and labor.

Task assignment and unit costs. At factor prices w_j and $1/q_j$, the unit cost of task x is:

$$c_j(x) = \min\left\{ \frac{w_j}{A_j^H z_j^H(x)}, \frac{1}{q_j A_j^K z_j^K(x)} \right\}. \tag{A.5}$$

The cheaper factor supplies the entire task; ties have measure zero. Define the following objects:

$$v_j^H = (A_j^H / w_j)^{\eta-1}, \qquad v_j^K = (A_j^K q_j)^{\eta-1}, \qquad \Psi_j = v_j^H + v_j^K. \tag{A.6}$$

The two candidate unit costs raised to the power $\eta - 1$ are independent exponential variables with rates v_j^H and v_j^K. The exponential-race formula then gives:

$$\Pr\{H \text{ supplies task } x\} = \frac{v_j^H}{\Psi_j}, \qquad \Pr\{c_j(x) > c\} = \exp(-\Psi_j c^{\eta-1}). \tag{A.7}$$

For example, the first probability equals $\int_0^\infty v_j^H \exp(-\Psi_j t)\,dt$. The cost of the unnormalized task bundle is $(\int_0^1 c_j(x)^{1-\sigma} dx)^{1/(1-\sigma)}$. Using the change of variables $u = \Psi_j c^{\eta-1}$ yields:

$$\begin{aligned} \mathbb{E}[c_j^{1-\sigma}] &= \Psi_j^{-(1-\sigma)/(\eta-1)} \int_0^\infty u^{(1-\sigma)/(\eta-1)} e^{-u}\,du \\ &= \Gamma\left(1 + \frac{1-\sigma}{\eta-1}\right) \Psi_j^{-(1-\sigma)/(\eta-1)}. \end{aligned} \tag{A.8}$$

The condition that $\sigma < \eta$ ensures that this moment is finite. Dividing the bundle's unit cost by the normalization in (A.4) gives:

$$p_j = \frac{\mathbb{E}[c_j^{1-\sigma}]^{1/(1-\sigma)}}{\kappa_{\sigma,\eta}} = \left[(A_j^H/w_j)^{\eta-1} + (A_j^K q_j)^{\eta-1}\right]^{-1/(\eta-1)}. \tag{A.9}$$

This is the dual unit-cost function of the occupation-level CES technology. Hence, the efficient occupation-level production frontier is:

$$y_j(h_j, k_j) = \left[(A_j^H h_j)^{(\eta-1)/\eta} + (A_j^K k_j)^{(\eta-1)/\eta}\right]^{\eta/(\eta-1)}. \tag{A.10}$$

Without the normalization, the right-hand side would instead be multiplied by $\kappa_{\sigma,\eta}^{-1}$. Equivalently this constant could be absorbed into both factor productivities.

Why factor cost shares equal task shares. The minimizing cost is independent of the identity of the cheaper factor. For any $c \geq 0$,

$$\Pr\{c_j(x) > c,\ H \text{ supplies } x\} = \int_{c^{\eta-1}}^{\infty} v_j^H \exp(-\Psi_j t)\, dt = \frac{v_j^H}{\Psi_j} \exp(-\Psi_j c^{\eta-1}). \tag{A.11}$$

The CES demand structure implies that expenditure on task x is proportional to $c_j(x)^{1-\sigma}$. Independence then implies that the labor share of expenditures equals the measure of labor-supplied tasks:

$$\begin{aligned} s_j^H &= \frac{w_j h_j}{p_j y_j} = \frac{\mathbb{E}[\mathbf{1}\{H \text{ supplies } x\} c_j(x)^{1-\sigma}]}{\mathbb{E}[c_j(x)^{1-\sigma}]} \\ &= \frac{(A_j^H/w_j)^{\eta-1}}{(A_j^H/w_j)^{\eta-1} + (A_j^K q_j)^{\eta-1}}, \qquad s_j^K = 1 - s_j^H. \end{aligned}$$

Consequently, s_j^K is also the share of tasks assigned to capital.

B.2 Derivation of the inverse labor demand curve (13)

The representative firm maximizes profits, taking factor prices as given:

$$\max_{h_j,k_j,y_j} \left(\sum_j \alpha_j^{1/\lambda} y_j^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}} - \sum_j (w_j h_j + k_j/q_j) + \sum_j p_j \left[\left((A_j^H h_j)^{\frac{\eta-1}{\eta}} + (A_j^K k_j)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} - y_j \right],$$

where $p_j > 0$. The corresponding first-order conditions are:

$$p_j = \left(\frac{\alpha_j Y}{y_j} \right)^{1/\lambda} \tag{A.12}$$

$$w_j = p_j \left(\frac{y_j}{h_j} \right)^{1/\eta} (A_j^H)^{\frac{\eta-1}{\eta}} \tag{A.13}$$

$$\frac{1}{q_j} = p_j \left(\frac{y_j}{k_j} \right)^{1/\eta} (A_j^K)^{\frac{\eta-1}{\eta}}. \tag{A.14}$$

Solving for k_j in (A.14) and substituting into the expression for y_j gives:

$$y_j = (1 - [A_j^K q_j p_j]^{\eta-1})^{\frac{\eta}{1-\eta}} \cdot A_j^H h_j.$$

Note that $p_j^{1-\eta} = (A_j^H/w_j)^{\eta-1} + (A_j^K q_j)^{\eta-1}$ and $(A_j^K q_j p_j)^{\eta-1} = s_j^K = 1 - s_j^H$. Thus, aggregate output can be written as:

$$Y = \left(\sum_j \alpha_j^{1/\lambda} \left((s_j^H)^{\frac{\eta}{1-\eta}} \cdot A_j^H \cdot h_j \right)^{\frac{\lambda-1}{\lambda}} \right)^{\frac{\lambda}{\lambda-1}}.$$

Solving for y_j in (A.12) and substituting into (A.13) then gives:

$$w_j = p_j^{\frac{\eta-\lambda}{\eta}} \left(\frac{\alpha_j Y}{h_j} \right)^{1/\eta} (A_j^H)^{\frac{\eta-1}{\eta}}.$$

Noting that $p_j = (s_j^H)^{\frac{1}{\eta-1}} (w_j/A_j^H)$ and evaluating h_j at $h_j(\boldsymbol{w})$ then gives the final result. When $\eta > 1$, in order for output to be finite, we thus require:

$$\left(\sum_j \alpha_j (A_j^K q_j)^{\lambda-1} \right)^{\frac{1}{1-\lambda}} > 1, \qquad \text{if } \lambda \neq 1,$$

$$\prod_j (A_j^K q_j)^{-\alpha_j} > 1, \qquad \text{if } \lambda = 1.$$

This condition resembles the necessary condition in Acemoglu and Restrepo (2022) that accounts for the roundabout nature of production.

B.3 Technology comparative statics

Let $\hat{x} \coloneqq d \log x$. Log-differentiating (13) gives:

$$\begin{aligned}
\hat{w}_j &= \xi \cdot \hat{s}_j^H + \frac{1}{\lambda} \cdot \hat{Y} - \frac{1}{\lambda} \cdot \hat{h}_j + \frac{\lambda - 1}{\lambda} \cdot \hat{A}_j^H \\
\Rightarrow \left(1 + \left(\frac{\eta - \lambda}{\lambda}\right) \cdot s_j^K\right) \hat{w}_j &= \frac{1}{\lambda} \cdot \hat{Y} - \left(\frac{\eta - \lambda}{\lambda}\right) \cdot s_j^K \cdot (\hat{A}_j^K + \hat{q}_j) + \left(\frac{(\eta - 1) - (\eta - \lambda) s_j^H}{\lambda}\right) \cdot \hat{A}_j^H - \frac{1}{\lambda} \cdot \hat{h}_j \\
\varepsilon_j^d \cdot \hat{w}_j &= \hat{Y} - (\eta - \lambda) \cdot s_j^K \cdot (\hat{A}_j^K + \hat{q}_j) + \left(\varepsilon_j^d - 1\right) \cdot \hat{A}_j^H - \hat{h}_j \\
&= \hat{Y} + z_j - \sum_k \varepsilon_{jk}^S \cdot \hat{w}_k \\
\Rightarrow \hat{w}_j &= \frac{\hat{Y} + z_j - \sum_{k \neq j} \varepsilon_{jk}^S \cdot \hat{w}_k}{\varepsilon_j^d + \varepsilon_{jj}^S},
\end{aligned}$$

where $\varepsilon_j^d \coloneqq \eta \cdot s_j^K + \lambda \cdot s_j^H > 0$ is the conditional elasticity of labor demand in occupation j with respect to wages in j (holding fixed technology and aggregate output) and $z_j = -(\eta - \lambda) \cdot s_j^K \cdot (\hat{A}_j^K + \hat{q}_j) + (\varepsilon_j^d - 1) \cdot \hat{A}_j^H$ is the direct effect of changes in technology on labor demand in occupation j. The second equality uses the fact that $\hat{s}_j^H = (\eta - 1) \cdot s_j^K \left(\hat{A}_j^H - \hat{w}_j - \hat{A}_j^K - \hat{q}_j\right)$, and the fourth equality uses the fact that $\hat{h}_j = \sum_k \varepsilon_{jk}^S \cdot \hat{w}_k$. When $\eta > \lambda$, the direct effect of improvements in capital efficiency in j on w_j is therefore negative. Similarly, note that $\varepsilon_j^d - 1 = (\eta - 1) s_j^K + (\lambda - 1) s_j^H$. Hence, the direct effect of labor-augmenting technological change in j on w_j is unambiguously positive whenever both $\eta > 1$ and $\lambda > 1$.

There are two instructive limiting cases. When occupations are perfect substitutes ($\lambda \to \infty$), only the direct effect remains and $\hat{w}_j = (\hat{A}_j^K + \hat{q}_j) \cdot s_j^K / s_j^H > 0$, since labor in j must benefit from improvements in technology with perfectly elastic demand at the occupation level and a perfectly elastic capital supply. In this case, the resulting change in automation in j, implied by Equation (17), is given by $\hat{s}_j^K = (\eta - 1)(\hat{A}_j^K + \hat{q}_j)$, which is larger than the change induced by the direct response. In contrast, as capital and labor become perfect substitutes in production in each occupation ($\eta \to \infty$), only the direct effect remains for the occupations with capital actively used in production, $\hat{w}_j = -(\hat{A}_j^K + \hat{q}_j)$.

Re-writing Equation (16) in matrix notation and rearranging, we obtain:

$$\hat{\boldsymbol{w}} = \boldsymbol{\Theta}^{-1} \left(\hat{Y} \cdot \mathbf{1} + \boldsymbol{z}\right), \text{ where } \boldsymbol{\Theta} = \boldsymbol{\varepsilon^D} + \boldsymbol{\varepsilon^S} \text{ and } \boldsymbol{\varepsilon^D} = \operatorname{diag}(\varepsilon_j^d).$$

As shown in Appendix B.4, the labor supply elasticity matrix has zero row sums, $\sum_k \varepsilon_{jk}^S = 0$ for every j. It then follows that:

$$\boldsymbol{\Theta} \cdot \mathbf{1} = \boldsymbol{\varepsilon^D} \cdot \mathbf{1} + \boldsymbol{\varepsilon^S} \cdot \mathbf{1} = \boldsymbol{\varepsilon^D} \cdot \mathbf{1}.$$

With common labor demand elasticities, $\varepsilon_j^d \equiv \varepsilon^d$ (when capital shares are initially constant across occupations, $s_j^K \equiv s^K$), it follows that:

$$\boldsymbol{\Theta}^{-1} \cdot \mathbf{1} = \frac{1}{\varepsilon^d} \cdot \mathbf{1} \Rightarrow \hat{\boldsymbol{w}} = \frac{\hat{Y}}{\varepsilon^d} \cdot \mathbf{1} + \boldsymbol{\Theta}^{-1} \cdot \boldsymbol{z} = \frac{\hat{Y}}{\varepsilon^d} \cdot \mathbf{1} + \left(\mathbf{I} + \frac{\boldsymbol{\varepsilon^S}}{\varepsilon^d}\right)^{-1} \cdot \frac{\boldsymbol{z}}{\varepsilon^d},$$

which is similar to the expression in Fan (2025). In this sense, $\boldsymbol{\Theta}^{-1} \cdot \boldsymbol{z}$ summarizes all relative changes in

real wages in response to uneven labor demand shocks: $[\boldsymbol{\Theta}^{-1} \cdot \boldsymbol{z}]_j = \hat{w}_j - \hat{Y}/\varepsilon^d$. Additionally, in this case, $\boldsymbol{\Theta}$ is invertible whenever $-\varepsilon^d$ is not an eigenvalue of $\boldsymbol{\varepsilon^S}$. With heterogeneous capital shares, however, the elements of $\boldsymbol{\Theta}^{-1} \cdot \mathbf{1}$ generally differ across occupations, so changes in aggregate output will also affect relative wages. The size of these wage responses depends jointly on occupation-specific labor demand elasticities and cross-occupation labor supply responses, provided $\boldsymbol{\Theta}$ is invertible. In general, with an unrestricted learning technology, we must verify numerically that $\boldsymbol{\Theta}$ is invertible at the estimated equilibrium.

Aggregate output. To derive the response of aggregate output, define the share of aggregate output spent on occupation j as $\pi_j = p_j y_j / Y$, with $\sum_j \pi_j = 1$, where all shares are evaluated at the initial equilibrium. Log-differentiating the ideal price condition gives $\sum_j \pi_j \cdot \hat{p}_j = 0$. From the occupation-level price changes, this implies that:

$$\sum_j \pi_j \cdot s_j^H \cdot \hat{w}_j = \sum_j \pi_j \left[s_j^K \cdot (\hat{A}_j^K + \hat{q}_j) + s_j^H \cdot \hat{A}_j^H \right].$$

Thus, the sum of skill-price changes weighted by each occupation's labor payments as a share of aggregate output equals the sum of technology changes weighted by their corresponding factor-cost shares in aggregate output. Define $\pi_j^H \coloneqq \pi_j \cdot s_j^H = (w_j h_j)/Y$ and $\pi_j^K \coloneqq \pi_j \cdot s_j^K = (k_j/q_j)/Y$, corresponding to the shares of labor and capital in occupation j in aggregate output. The log-differentiated ideal price condition can then be rewritten in vector notation as:

$$(\boldsymbol{\pi}^H)^\top \hat{\boldsymbol{w}} = (\boldsymbol{\pi}^H)^\top \hat{\boldsymbol{A}}^{\boldsymbol{H}} + (\boldsymbol{\pi}^K)^\top (\hat{\boldsymbol{A}}^{\boldsymbol{K}} + \hat{\boldsymbol{q}}).$$

Substituting the wage response derived above and collecting terms in $\hat{Y}$:

$$(\boldsymbol{\pi}^H)^\top (\boldsymbol{\Theta}^{-1} \cdot \hat{Y} \cdot \mathbf{1}) + (\boldsymbol{\pi}^H)^\top (\boldsymbol{\Theta}^{-1} \cdot \boldsymbol{z}) = (\boldsymbol{\pi}^H)^\top \hat{\boldsymbol{A}}^{\boldsymbol{H}} + (\boldsymbol{\pi}^K)^\top (\hat{\boldsymbol{A}}^{\boldsymbol{K}} + \hat{\boldsymbol{q}}).$$

Provided $(\boldsymbol{\pi}^H)^\top (\boldsymbol{\Theta}^{-1} \cdot \mathbf{1}) \neq 0$, the aggregate output response is therefore:

$$\hat{Y} = \frac{(\boldsymbol{\pi}^H)^\top \hat{\boldsymbol{A}}^{\boldsymbol{H}} + (\boldsymbol{\pi}^K)^\top (\hat{\boldsymbol{A}}^{\boldsymbol{K}} + \hat{\boldsymbol{q}}) - (\boldsymbol{\pi}^H)^\top (\boldsymbol{\Theta}^{-1} \cdot \boldsymbol{z})}{(\boldsymbol{\pi}^H)^\top (\boldsymbol{\Theta}^{-1} \cdot \mathbf{1})}, \tag{A.15}$$

which requires that $(\boldsymbol{\pi}^H)^\top (\boldsymbol{\Theta}^{-1} \cdot \mathbf{1}) \neq 0$. With common labor demand elasticities, $\boldsymbol{\Theta}^{-1} \cdot \mathbf{1} = (1/\varepsilon^d) \cdot \mathbf{1}$, so expression (A.15) simplifies to:

$$\begin{aligned} \hat{Y} &= \frac{\varepsilon^d}{\sum_j \pi_j^H} \left[(\boldsymbol{\pi}^H)^\top \hat{\boldsymbol{A}}^{\boldsymbol{H}} + (\boldsymbol{\pi}^K)^\top (\hat{\boldsymbol{A}}^{\boldsymbol{K}} + \hat{\boldsymbol{q}}) - (\boldsymbol{\pi}^H)^\top (\boldsymbol{\Theta}^{-1} \cdot \boldsymbol{z}) \right] \\ &= \sum_j \pi_j \hat{A}_j^H + \frac{\eta s^K}{s^H} \sum_j \pi_j (\hat{A}_j^K + \hat{q}_j) + \boldsymbol{\pi}^T \cdot \boldsymbol{\varepsilon}^S \cdot \boldsymbol{\Theta}^{-1} \boldsymbol{z}, \end{aligned}$$

where the third term depends on the resulting change in effective labor supply. Next, we can solve for the response in occupation-level capital shares. Since $s_j^K = 1 - s_j^H$, $ds_j^K = -ds_j^H$ and $s_j^K \hat{s}_j^K = -s_j^H \hat{s}_j^H$. The

proportional change in the capital share is therefore given by:

$$\begin{aligned}\hat{s}_j^K &= (\eta-1)\cdot s_j^H \cdot \left(\hat{w}_j + \hat{A}_j^K + \hat{q}_j - \hat{A}_j^H\right) \\ &= (\eta-1)\cdot s_j^H \cdot \left(\left[\boldsymbol{\Theta}^{-1}\left(\hat{Y}\cdot\mathbf{1}+\boldsymbol{z}\right)\right]_j + \hat{A}_j^K + \hat{q}_j - \hat{A}_j^H\right),\end{aligned}$$

where $\hat{Y}$ is given by (A.15). Note that even if occupation j experiences no direct change in technology, its capital share will still respond to movements in equilibrium wages:

$$\hat{s}_j^K = (\eta-1)\cdot s_j^H \cdot [\boldsymbol{\Theta}^{-1}\cdot(\hat{Y}\cdot\mathbf{1}+\boldsymbol{z})]_j.$$

If $\varepsilon_j^d \equiv \varepsilon^d$, then $\hat{s}_j^K = (\eta-1)\cdot s^H\cdot(\hat{Y}/\varepsilon^d + [\boldsymbol{\Theta}^{-1}\cdot\boldsymbol{z}]_j)$. This last equation illustrates how the reallocation of labor supply then directly determines the *differential* response of capital shares in occupations that are not directly affected by the automation shock.

Aggregate capital share. Define the aggregate capital share as S^K. Its response is given by:

$$\begin{aligned}dS^K = &\underbrace{\sum_j \pi_j ds_j^K}_{\text{within occupations}} + \underbrace{\sum_j \pi_j s_j^K \hat{\pi}_j}_{\text{expenditure reweighting}} \\ &= (\eta-1)\sum_j \pi_j s_j^K(\hat{A}_j^K + \hat{q}_j) + (\eta-\lambda)\sum_j \pi_j s_j^K \hat{p}_j. \qquad \text{(A.16)}\end{aligned}$$

Equivalently, the second term is $(\eta-\lambda)\operatorname{Cov}_\pi(s^K,\hat{p})$, where the covariance is taken under the initial expenditure weights. To see this, note that (A.16) can be re-written as:

$$\begin{aligned}dS^K &= \sum_j \pi_j(\eta-1)s_j^K(\hat{p}_j + \hat{A}_j^K + \hat{q}_j) + \sum_j \pi_j s_j^K(1-\lambda)\hat{p}_j \\ &= (\eta-1)\sum_j \pi_j s_j^K(\hat{A}_j^K + \hat{q}_j) + (\eta-\lambda)\sum_j \pi_j s_j^K \hat{p}_j.\end{aligned}$$

Because $\sum_j \pi_j\cdot\hat{p}_j = 0$, the second term can be re-written as: $\sum_j \pi_j s_j^K\hat{p}_j = \sum_j \pi_j(s_j^K - S^K)\hat{p}_j = \operatorname{Cov}_\pi(s^K,\hat{p})$. To express (A.16) in terms of technology shocks and the labor supply elasticity matrix, use the following:

$$\hat{p}_j = s_j^H\left(\left[\boldsymbol{\Theta}^{-1}(\hat{Y}\mathbf{1}+\boldsymbol{z})\right]_j - \hat{A}_j^H\right) - s_j^K(\hat{A}_j^K + \hat{q}_j),$$

with $\hat{Y}$ from (A.15). Finally, note that if $s_j^K \equiv s^K$ at the initial equilibrium, then for any $\lambda > 0$:

$$dS^K = (\eta-1)s^K\sum_j \pi_j(\hat{A}_j^K + \hat{q}_j),$$

which implies that the first-order aggregate capital-share response is independent of the labor-supply elasticity matrix, *conditional on the same initial expenditure shares* and technology shock.

B.4 Labor supply responses

Employment responses. Differentiating the old workers' logit probabilities and inclusive value gives:

$$\frac{\partial \log e_j^{o|\ell}}{\partial \log w_k} = \frac{\mathbb{1}[j=k] - e_k^{o|\ell}}{\nu}, \qquad \frac{\partial \bar{V}^{o|\ell}}{\partial \log w_k} = e_k^{o|\ell}. \tag{A.17}$$

Differentiating the young workers' logit probabilities therefore gives:

$$\frac{\partial \log e_j^{y}}{\partial \log w_k} = \frac{1}{\nu}\left[\mathbb{1}[j=k] - e_k^{y} + \beta\left(e_k^{o|j} - \sum_\ell e_\ell^{y} e_k^{o|\ell}\right)\right]. \tag{A.18}$$

For $j \neq k$, young employment in j increases with w_k exactly when:

$$\beta\left(e_k^{o|j} - \sum_\ell e_\ell^{y} e_k^{o|\ell}\right) > e_k^{y}. \tag{A.19}$$

This compares the advantage of entering j first and reaching k when old with the attraction of working directly in k when young. For the own-wage response, rewrite the continuation term as $e_j^{o|j} - \sum_\ell e_\ell^{y} e_j^{o|\ell} = \sum_{\ell \neq j} e_\ell^{y}(e_j^{o|j} - e_j^{o|\ell})$. Since each difference lies in $[-1, 1]$, it follows that:

$$0 < \frac{(1-\beta)(1-e_j^{y})}{\nu} \leq \frac{\partial \log e_j^{y}}{\partial \log w_j} \leq \frac{(1+\beta)(1-e_j^{y})}{\nu}. \tag{A.20}$$

Strict positivity uses interior choices with $J \geq 2$ and $\beta < 1$. This result concerns the labor supply response of young workers. It does not by itself sign the aggregate effective-labor response, which also includes the dynamic composition effect. In practice, I always find strictly positive own-wage responses, reflecting the strong patterns of occupational persistence in the data.

Labor supply. For independent unit-scale Fréchet shocks of shape $1/\nu$, the conditional mean of the selected productivity draw in an alternative chosen with probability e is $\Gamma(1-\nu)e^{-\nu}$, which gives the expressions for the effective labor supply of young and old workers in equations (10) and (11):

$$\begin{aligned} h_j^{y} &= \Gamma(1-\nu)\psi_j^{y}(e_j^{y})^{1-\nu}, \\ h_j^{o|\ell} &= \Gamma(1-\nu)\psi_j^{y}\gamma_{\ell j} e_\ell^{y}(e_j^{o|\ell})^{1-\nu}. \end{aligned}$$

Log-differentiating each contribution to (9) and then summing gives the expression in (18):

$$\varepsilon_{jk}^{S} = \underbrace{(1-\nu)\frac{h_j^{y}}{h_j}\frac{\partial \log e_j^{y}}{\partial \log w_k}}_{\text{Young reallocation}} + \underbrace{(1-\nu)\sum_\ell \frac{h_j^{o|\ell}}{h_j}\frac{\partial \log e_j^{o|\ell}}{\partial \log w_k}}_{\text{Old reallocation}} + \underbrace{\sum_\ell \frac{h_j^{o|\ell}}{h_j}\frac{\partial \log e_\ell^{y}}{\partial \log w_k}}_{\text{Dynamic composition}}.$$

Selection attenuates the first two components because the relevant employment probabilities enter effective labor with exponent $1-\nu$. The composition term has no such multiplier: the origin mass e_ℓ^{y} enters old effective labor linearly, and old productivity shocks are independent of young shocks.

The dynamic composition effect. Because $\sum_\ell e^y_\ell = 1$,

$$\sum_\ell e^y_\ell \cdot \frac{\partial \log e^y_\ell}{\partial \log w_k} = 0. \tag{A.21}$$

Substitute (11) into the composition term in (18) and use (A.21):

$$\sum_\ell \frac{h^{o|\ell}_j}{h_j} \frac{\partial \log e^y_\ell}{\partial \log w_k} = \frac{\Gamma(1-\nu)\psi^y_j}{h_j} \operatorname{Cov}_{e^y_\ell}\left(\gamma_{\ell j}(e^{o|\ell}_j)^{1-\nu}, \frac{\partial \log e^y_\ell}{\partial \log w_k}\right), \tag{A.22}$$

where $\operatorname{Cov}_{e^y_\ell}(x_\ell, z_\ell) = \sum_\ell e^y_\ell x_\ell z_\ell - (\sum_\ell e^y_\ell x_\ell)(\sum_\ell e^y_\ell z_\ell)$. The composition effect is positive when w_k reallocates young workers toward origins that generate more old effective labor supply in j, accounting for both the direct effect on human capital accumulation and subsequent occupational selection.

Log additive separability in γ. Suppose $\gamma_{\ell j} = g_\ell g_j$. The common origin factor g_ℓ cancels from old choice probabilities, so that we can re-write the old choice probabilities as:

$$e^{o|\ell}_j = e^o_j = \frac{(w_j \psi^y_j g_j)^{1/\nu} \exp(a_j/\nu)}{\sum_k (w_k \psi^y_k g_k)^{1/\nu} \exp(a_k/\nu)}. \tag{A.23}$$

Thus, the continuation term in (A.18) vanishes and $\partial \log e^y_\ell / \partial \log w_k = [\mathbb{1}[\ell = k] - e^y_k]/\nu$. Learning still affects young employment *levels*, since $\bar{V}^{o|\ell} = \log g_\ell + \bar{V}^o(\boldsymbol{w}, \boldsymbol{g} \odot \boldsymbol{\psi}^y)$. Define $\bar{g} = \sum_\ell e^y_\ell g_\ell$ and $h^o_j = \sum_\ell h^{o|\ell}_j$. Then, we can re-write old worker effective labor supply in j as:

$$h^o_j = \Gamma(1-\nu)\psi^y_j g_j (e^o_j)^{1-\nu} \bar{g}, \qquad \frac{\partial \log \bar{g}}{\partial \log w_k} = \frac{e^y_k}{\nu}\left(\frac{g_k}{\bar{g}} - 1\right). \tag{A.24}$$

Holding old choice probabilities and their associated selection fixed isolates the composition effect:

$$\sum_\ell \frac{h^{o|\ell}_j}{h_j} \frac{\partial \log e^y_\ell}{\partial \log w_k} = \frac{h^o_j}{h_j} \frac{e^y_k}{\nu}\left(\frac{g_k}{\bar{g}} - 1\right). \tag{A.25}$$

The old labor share h^o_j/h_j converts its composition response into a total-labor response. An increase in w_k improves old labor's productive composition everywhere when $g_k > \bar{g}$, despite origin-independent old choices. The effect vanishes when g_ℓ is constant across origins.

Homogeneity. Summing (A.17) or (A.18) over wage indices k gives zero (a product of log utility). Applying the same summation to (18) therefore gives:

$$\sum_k \varepsilon^S_{jk} = 0 \quad \text{for every } j, \qquad \boldsymbol{\varepsilon}^S \cdot \mathbf{1} = \mathbf{0}. \tag{A.26}$$

A common proportional wage increase does not induce reallocation or a change in effective labor.

B.5 Toy example details

Proof of the wage responses. Write $\hat{x} = d\log x$. Differentiating career indifference and subtracting labor demand in the two occupations along the training career gives:

$$(\hat{w}_2 - \hat{w}_1) + \beta(\hat{w}_3 - \hat{w}_1) = 0, \tag{A.27}$$

$$\varepsilon^d(\hat{w}_2 - \hat{w}_3) = -(\eta - \lambda)s^K(\mathbb{1}[i=2] - \mathbb{1}[i=3])\tau. \tag{A.28}$$

The second equality uses $h_2 = h_3 = e_2^y$ and the labor-demand equations:

$$\varepsilon^d \hat{w}_j = \hat{Y} - (\eta - \lambda)s^K \mathbb{1}[j=i]\tau - \hat{h}_j. \tag{A.29}$$

Solving the first two equations gives Δ_i and the relative wage responses. The final-good price index implies:

$$\hat{p}_j = s^H \hat{w}_j - s^K \mathbb{1}[j=i]\tau, \qquad \sum_j \pi_j \cdot \hat{p}_j = 0, \qquad \sum_j \pi_j \cdot \hat{w}_j = \frac{s^K}{s^H}\pi_i \tau.$$

Substituting the relative wage responses gives the first wage equation in the remark. Since $\pi_3/\pi_2 = w_3/w_2 > 1$, we have $\pi_3 - \beta\pi_2 > 0$. Under training automation, this implies that $\hat{w}_3 > \hat{w}_1 > \hat{w}_2$ and the positive weighted mean wage response implies $\hat{w}_3 > 0$. Under destination automation, $\hat{w}_2 > \hat{w}_1 > \hat{w}_3$ with $\hat{w}_1 > 0$. For either shock along the training path, the shocked occupation's capital-share response exceeds the linked occupation's response by:

$$(\eta - 1)s^H\left[\tau - (1+\beta)\Delta_2\right] = \frac{(\eta-1)s^H\lambda}{\varepsilon^d}\tau > 0.$$

The remaining rankings follow directly from relative wages. Importantly, these are rankings of proportional wage changes, rather than wage levels. Since initial capital shares are common, changes in capital shares have the same ranking.

Training entry and aggregates. Subtract occupation one's labor-demand equation from occupation three's and use $\hat{h}_1 = -de_2^y/(1-e_2^y)$ and $\hat{h}_3 = de_2^y/e_2^y$. For a shock to occupation i, it follows that:

$$\frac{de_2^{y,(i)}}{e_2^y(1-e_2^y)} = (\eta - \lambda)s^K(\mathbb{1}[i=1] - \mathbb{1}[i=3])\tau - \varepsilon^d\Delta_i.$$

Consequently, changes in entry into the career path are given by:

$$\begin{aligned} de_2^{y,(1)} &= (\eta-\lambda)s^K e_2^y(1-e_2^y)\tau > 0, \\ de_2^{y,(2)} &= -\frac{(\eta-\lambda)s^K e_2^y(1-e_2^y)}{1+\beta}\tau < 0, \\ de_2^{y,(3)} &= \beta\, de_2^{y,(2)} < 0. \end{aligned} \tag{A.30}$$

For the aggregate capital share, initially common factor shares imply $\sum_j s^K \cdot d\pi_j = 0$. Thus

$$dS^K = s^K \sum_j \pi_j \hat{s}_j^K = (\eta - 1)s^K \pi_i \tau > 0. \tag{A.31}$$

Weighting (A.29) by π_j and using $\pi_j = w_j h_j/(s^H Y)$ gives the gross-output response:

$$\hat{Y}^{(i)} = \eta \frac{s^K}{s^H} \pi_i \tau + \frac{w_2 + w_3 - 2w_1}{s^H Y} \, de_2^{y,(i)}. \tag{A.32}$$

The composition coefficient is positive: career indifference gives $w_2/w_1 = (w_3/w_1)^{-\beta}$, and $w_3/w_1 + (w_3/w_1)^{-\beta} > 2$ for $w_3/w_1 > 1$ and $0 < \beta < 1$. Outside automation therefore raises output. Destination automation also raises output because its outside wage response is positive and (A.29) implies $\hat{Y}^{(3)} = \varepsilon^d \hat{w}_1^{(3)} - de_2^{y,(3)}/(1-e_2^y) > 0$. For training automation, the sign of the output response is generally ambiguous:

$$\hat{Y}^{(2)} = \varepsilon^d \left[\frac{s^K}{s^H} \pi_2 \tau + (e_2^y - \pi_3 + \beta \pi_2) \Delta_2 \right].$$

This is positive for wage ratios sufficiently close to one, but negative for sufficiently large w_3/w_1.

Technology primitives. Choose e_2^y, s^K, and $s \in (0,1)$. The below construction targets $w_1 = 1$, $1 < w_3 < 1/s$, and $w_2 = w_3^{-\beta}$, which satisfy the assumptions in Section 4.2. With $h_1 = 2(1 - e_2^y)$ and $h_2 = h_3 = e_2^y$, choose production primitives:

$$\alpha_j = \frac{w_j h_j}{\sum_k w_k h_k}, \qquad A_j^H = w_j (s^H)^{1/(\eta-1)}, \qquad A_j^K q_j = (s^K)^{1/(\eta-1)}. \tag{A.33}$$

Any positive q_j can be chosen, with A_j^K set accordingly. The CES unit cost then equals one, with factor shares s^H, s^K. The final-good price index is also one, including at $\lambda = 1$. Setting $Y = \sum_j w_j h_j / s^H$, $y_j = \alpha_j Y$, and $k_j/q_j = s^K y_j$ satisfies all production first-order conditions and market-clearing equations. The capital-only final-good unit cost is $(s^K)^{-1/(\eta-1)} > 1$, so capital alone cannot generate unbounded output.

C Estimation Appendix (Section 5)

C.1 Data construction for estimation of the labor supply parameters

ACS sample construction and data details. I restrict the ACS sample to civilian wage and salary workers who worked at least 14 weeks in the preceding year, with implied annual hours between 250 and 4,500 and annual wage and salary earnings of at least $1,000 in 2012 dollars. Results are robust to additionally requiring full-time (30+ hours), full-year (48+ weeks) employment. I exclude self-employed workers, unpaid family workers, and individuals residing in institutional group quarters. I construct annual hours as usual weekly hours multiplied by the midpoint of the reported weeks-worked interval. Annual wage and salary earnings in the ACS are top-coded at the state level during this period, with earnings in the top 0.5% of the state-specific distribution replaced by the mean earnings of individuals above the corresponding threshold. I do not make further adjustments for top-coded earnings, given the small fraction of observations affected in the estimation sample.

To construct type- and age-specific occupational employment shares and average log earnings, I first compute these moments separately by year and then average them across years. Before averaging mean log earnings, I remove aggregate wage growth using an annual wage index constructed from the ACS that holds fixed the educational and occupational composition of the sample. Finally, I re-normalize occupational employment shares to sum to one within each worker type.

Occupations. I aggregate ACS-SOC codes to the analogue of the broad (5-digit level) SOC codes, which yields 391 occupations. I further aggregate these broad ACS-SOC 5-digit occupations to obtain at least 100 observations in the pooled ACS data for each bachelor's-degree type in each of the 22–31 and 32–41 age groups, as follows. In a first step, I manually consolidate non-professional occupations (SOC majors 31-53) into SOC minor or major groups. In a second step, I combine occupations that remain below this coverage threshold with similar occupations within the same SOC major group, using O*NET-based measures of knowledge, skills, task content, and educational requirements to determine occupational similarity. The final list of occupations is reported in Table A.2.

Table A.2: Occupation categories

Analysis occupation	SOC level
Top executives	Whole minor (11-1)
Marketing and sales managers	Whole minor (11-2)
Operations, facilities, purchasing and logistics managers	Sub-minor (11-3)
Computer and information systems managers	Sub-minor (11-3)
Financial managers	Sub-minor (11-3)
Human resources managers	Sub-minor (11-3)
Other managers	Sub-minor (11-9)
Engineering, science, and construction managers	Sub-minor (11-9)
Education and childcare administrators	Sub-minor (11-9)
Hospitality managers	Sub-minor (11-9)
Medical, health, and social service managers	Sub-minor (11-9)

Table A.2: Occupation categories

Analysis occupation	SOC level
Property, real estate, and community association managers	Sub-minor (11-9)
Buyers, event planners, fundraisers, and talent agents	Sub-minor (13-1)
Claims adjusters, appraisers, examiners, and investigators	Sub-minor (13-1)
Compliance officers	Sub-minor (13-1)
Management analysts and business operations specialists	Sub-minor (13-1)
Human resources workers and specialists	Sub-minor (13-1)
Logisticians and project management specialists	Sub-minor (13-1)
Market research analysts and marketing specialists	Sub-minor (13-1)
Financial analysts, advisors, loan officers and appraisers	Sub-minor (13-2)
Accountants, auditors, and financial examiners	Sub-minor (13-2)
Computer occupations, except support	Sub-minor (15-1)
Computer support specialists	Sub-minor (15-1)
Mathematical science occupations	Whole minor (15-2)
Engineers and architects	Sub-minor (17-1, 17-2)
Surveyors, drafters and engineering technicians	Sub-minor (17-1, 17-3)
Life scientists	Whole minor (19-1)
Physical scientists	Whole minor (19-2)
Social scientists and related workers	Whole minor (19-3)
Science technicians and safety specialists	Whole minor (19-4, 19-5)
Counselors and social workers	Whole minor (21-1)
Religious workers	Whole minor (21-2)
Lawyers, judges, and related workers	Whole minor (23-1)
Legal support workers	Whole minor (23-2)
Postsecondary teachers	Whole minor (25-1)
Classroom teachers and instructors	Whole minor (25-2, 25-3)
Library, museum, and other education workers	Sub-minor (25-4, 25-9)
Teaching assistants	Sub-minor (25-9)
Performers, athletes and media production workers	Whole minor (27-2, 27-4)
Art and design workers	Whole minor (27-1)
Media and communication workers	Whole minor (27-3)
Health technologists and technicians	Whole minor (29-2)
Physicians and other doctoral clinicians	Sub-minor (29-1)
Nurses, therapists, and other practitioners	Sub-minor (29-1, 29-9)
Healthcare support occupations	Whole major (31)
Protective service occupations	Whole major (33)
Food preparation and serving related occupations	Whole major (35)
Building and grounds maintenance	Whole major (37)
Personal care and service occupations	Whole major (39)
Supervisors of sales workers	Whole minor (41-1)

Table A.2: Occupation categories

Analysis occupation	SOC level
Retail sales workers	Whole minor (41-2)
Sales representatives, services	Whole minor (41-3)
Sales representatives, wholesale and manufacturing	Whole minor (41-4)
Other sales and related workers	Whole minor (41-9)
Office supervisors, secretaries and administrative assistants	Whole minor (43-1, 43-6)
Other office support workers	Whole minor (43-2, 43-9)
Financial clerks	Whole minor (43-3)
Information and record clerks	Whole minor (43-4)
Material recording and dispatching workers	Whole minor (43-5)
Farming, fishing, and forestry occupations	Whole major (45)
Construction and extraction occupations	Whole major (47)
Installation, maintenance, and repair occupations	Whole major (49)
Production occupations	Whole major (51)
Transportation and material moving occupations	Whole major (53)

Notes: Table shows the occupational classification that I construct.

Major categories. I cluster majors using a sample-size-weighted k-means clustering algorithm, which minimizes squared Euclidean distances between majors' square-root occupational employment-share vectors and their cluster centroids. The square-root transformation makes the clustering less dominated by differences in large occupational shares.[74] Employment shares are measured among ACS graduates aged 22–31, and each group is required to contain at least 18% of sampled graduates to avoid types with limited empirical support. I use an LLM to assign descriptive labels to the clusters. Table A.3 lists the detailed majors assigned to each bachelor type. Figure A.9 shows how the different major groups differ in employment allocation across occupations and earnings within occupations, at ages 22–31.

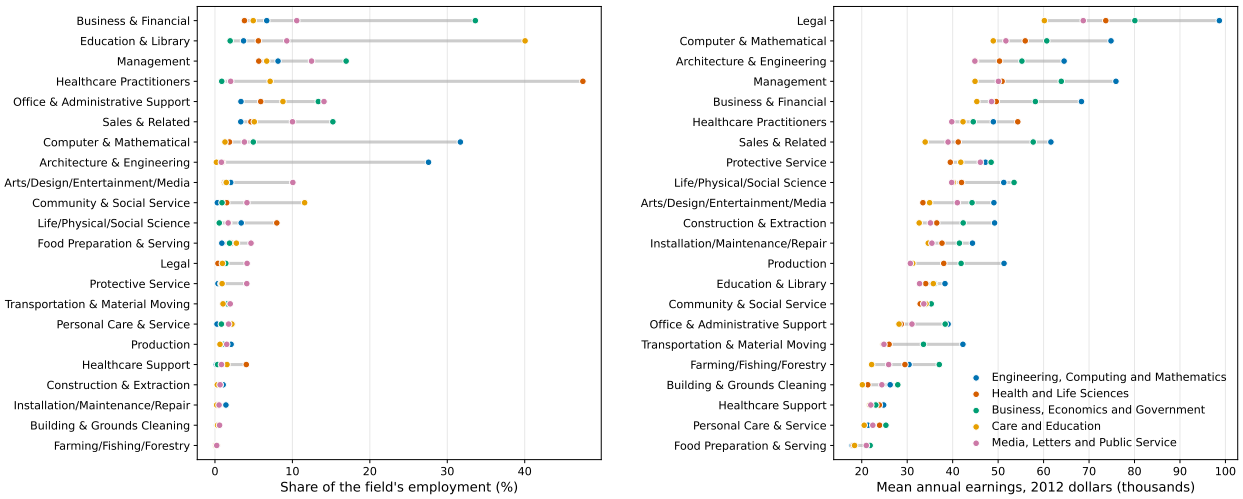
[74] Euclidean distances between the transformed share vectors are proportional to Hellinger distances between the corresponding employment distributions.

Table A.3: College major categories

Major group (% of graduates)	Largest fields (% of group)
Engineering, Computing and Mathematics (18%)	Computer Science 16, Mechanical Engineering 12, Electrical Engineering 8, Mathematics 6, Civil Engineering 6
Health and Life Sciences (18%)	Nursing 26, Biology 21, Physical Fitness Parks Recreation and Leisure 13, Chemistry 4, General Medical and Health Services 4
Business, Economics and Government (19%)	Business Management and Administration 23, Accounting 16, General Business 15, Marketing and Marketing Research 14, Finance 14
Care and Education (18%)	Psychology 30, Elementary Education 11, General Education 10, Social Work 6, Family and Consumer Sciences 5
Media, Letters and Public Service (26%)	Communications 11, Criminal Justice and Fire Protection 9, English Language and Literature 8, Political Science and Government 8, History 6

Notes: Table reports the largest detailed fields (majors) assigned to each type.

Figure A.9: Occupational employment and earnings by college major group



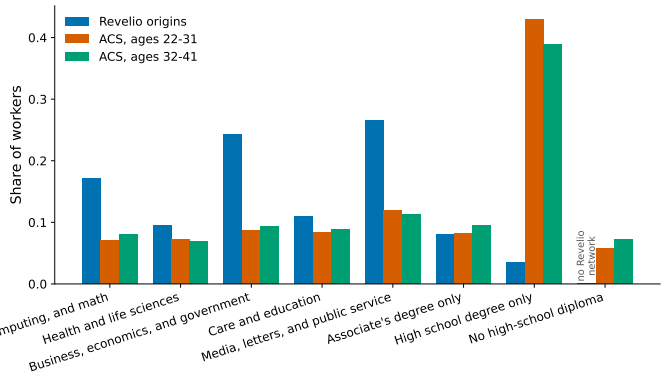
(a) Occupational employment shares

(b) Mean annual earnings within occupations

Notes: Figure uses pooled 2010–2022 ACS data. Panel (a) reports the share of each college major group's employment in each SOC major occupation group. Panel (b) reports mean annual earnings within each occupation and college major group, in thousands of 2012 dollars. Both panels report outcomes for workers aged 22–31. Colors identify the five college major groups in Table A.3; gray lines span the range across groups.

Distribution of workers across types. Figure A.10 shows the distribution of workers across types in Revelio and the ACS aged 22-31 and 32-41, before any normalization.

Figure A.10: Distribution of workers across types



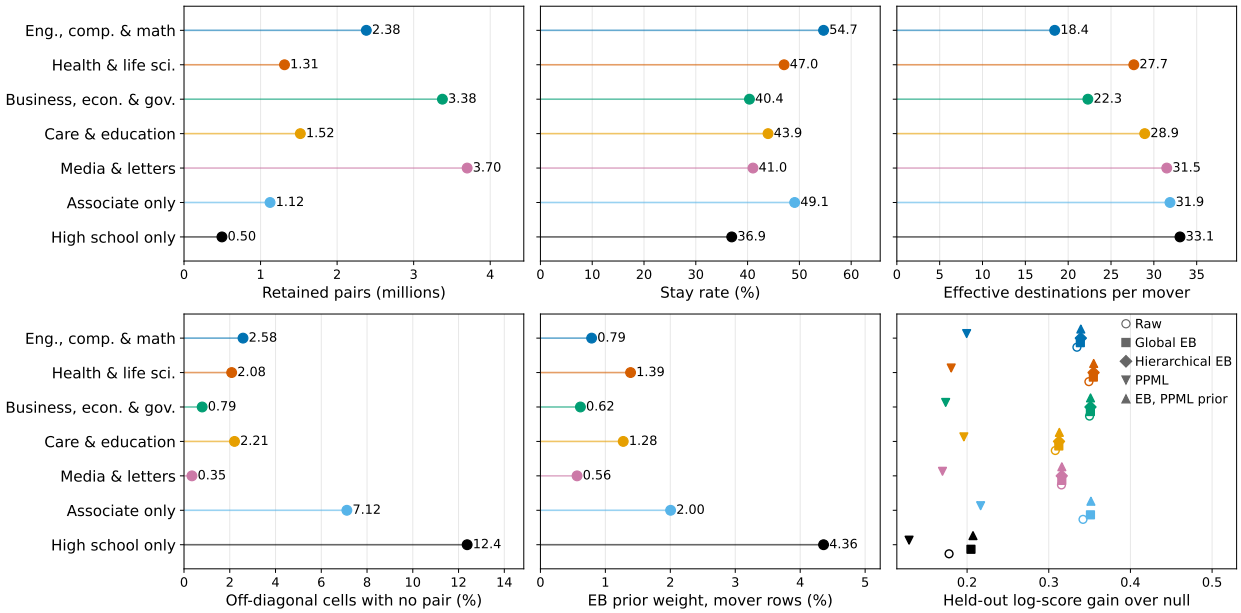
Notes: Figure shows the distribution of workers across types in Revelio and the ACS.

C.2 Constructing 10-year occupation transition networks

Empirical Bayes smoothing. To address sampling zeros and imprecisely measured transitions, I smooth the empirical transition networks using an empirical Bayes procedure that places a Beta prior on the probability of remaining in the initial occupation and a Dirichlet prior on destination probabilities conditional on moving. The prior strengths are estimated by maximizing the Beta-binomial and Dirichlet-multinomial marginal likelihoods across origins within each worker type. Because workers can contribute overlapping ten-year transitions from multiple origin years, I adjust the effective sample size used in the empirical Bayes procedure. Within each type–origin row, I divide all transition counts by the average number of observations per distinct worker. This preserves the empirical transition frequencies while basing the strength of regularization on the number of distinct workers rather than treating repeated observations as independent. However, the transition networks are very similar if I instead take the modal occupation of each user in each period (2010-2012 and 2020-2022).

Adjustment to match ACS shares. I use iterative proportional fitting (IPF), a procedure that alternately rescales origin–destination flows so that their row and column totals exactly match the ACS young and old occupational employment shares within each worker type. The resulting networks are very similar to those obtained by simply adjusting for occupational coverage differences in Revelio (the correlation across the two networks, pooled across types, is above 0.95). The coverage-based alternative re-weights destination probabilities by the ratio of ACS to Revelio cross-sectional employment shares for each occupation among workers aged 32–41 within each type before re-normalizing each origin's transition probabilities to sum to one. This correction excludes the group of workers without a high school degree, for whom I do not directly construct an empirical occupational transition network.

Figure A.11: Summary of transition networks by worker type



Notes: Figure summarizes the ten-year occupational transition networks for the seven worker types for which networks are estimated. The top row reports the number of retained observation pairs (not distinct workers), the share with the same occupation at both endpoints, and the effective number of destinations conditional on moving. The latter is the exponential of each origin's raw destination-share entropy, averaged across origins using mover counts as weights. The bottom row reports the percentage of off-diagonal cells with no observed pair, the global empirical Bayes prior weight on conditional moving probabilities (averaged using origin pair counts), and held-out predictive performance. The final panel compares raw frequencies, global EB, hierarchical EB, a PPML gravity predictor, and EB centered on the PPML-predicted conditional mover shares. The PPML specification includes origin and destination fixed effects, knowledge, skill, and task distances, industry overlap, and an indicator for membership in the same SOC major group. Workers are split into two halves, keeping all observations from each worker together. Each predictor is fitted on one half and evaluated on mover pairs in the other; the reported gains average across both directions. The null uses the training half's type-specific destination shares among movers, excluding each origin's own occupation and renormalizing. Scores are average log-probability gains over this null. All predictors, including the null, receive a 10^{-4} mixture weight on a uniform distribution over alternative destinations to keep log scores finite. Hierarchical estimates are reported only for the five bachelor's-degree types.

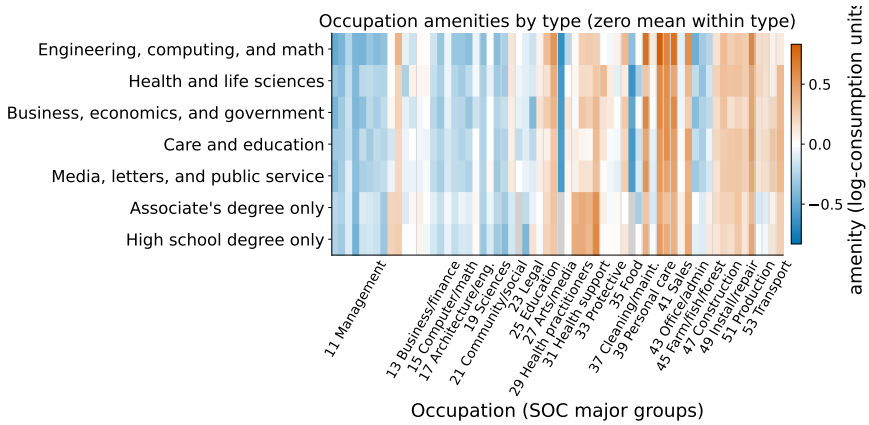
C.3 Additional details on the identification of labor supply parameters

Inverse shape parameter. I calibrate the inverse Fréchet shape parameter ν using the variance of log hourly wages within type–age group–occupation cells in the ACS. I first compute these variances separately by year and then pool them across years using effective-sample-size weights. I retain annual cells with a positive variance and an effective sample size of at least 30, and require at least eight usable years for each type–age group–occupation cell. For older workers, the observed variance of log wages additionally reflects differences in mean log wages among workers entering each destination from different origin occupations. I therefore subtract the model-implied component of wage dispersion arising from these differences. In the absence of additional measurement error, the remaining variance is given by $\pi^2\nu^2/6$.

Averaging these moments across types and occupations yields implied values of approximately $\nu = 0.39$ for young workers (ages 22–31) and $\nu = 0.41$ for older workers (ages 32–41). More generally, I find that the implied values of ν are similar across age groups within types, as well as across types, motivating the common benchmark value of $\nu = 0.4$, equivalent to a Fréchet shape parameter of 2.5. As an alternative, maximum likelihood estimates that use the full wage distribution yield a somewhat higher average value of approximately $\nu = 0.6$. Neither estimation approach explicitly accounts for measurement error in observed earnings, potentially leading to an upward bias in the estimated dispersion of idiosyncratic productivity shocks. Disentangling these two sources of wage dispersion would require additional assumptions on the distribution of measurement error.

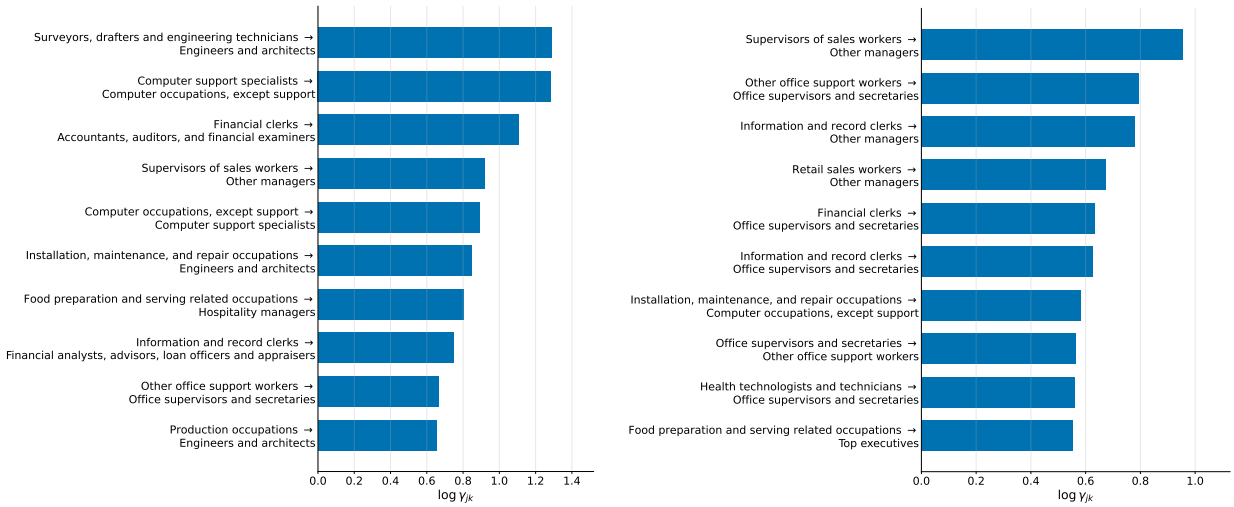
C.4 Additional figures and tables for Section 5

Figure A.12: Inferred occupational amenities by worker type



Notes: Figure reports inferred occupation-specific amenities for each worker type, measured in log-consumption units and normalized to have mean zero across occupations within each type. Columns correspond to occupations, grouped by SOC major category. Gray cells denote occupations that are closed to a given type due to educational requirements.

Figure A.13: Largest inferred learning gains: non-bachelor types

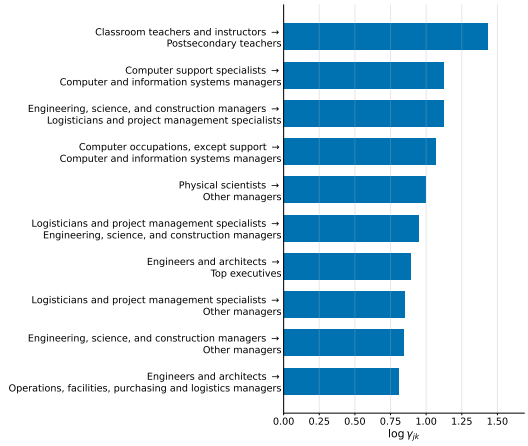


(a) Associate's degree only

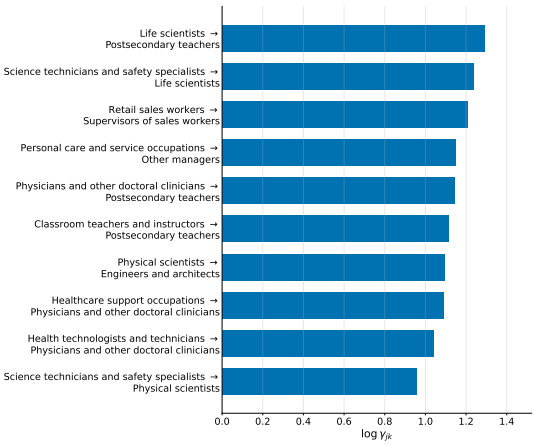
(b) High school degree only

Notes: Each panel reports the ten eligible off-diagonal occupation pairs with the largest inferred $\log \gamma_{mjk}$ for the indicated worker type. The ranking and eligibility criteria follow Figure A.14.

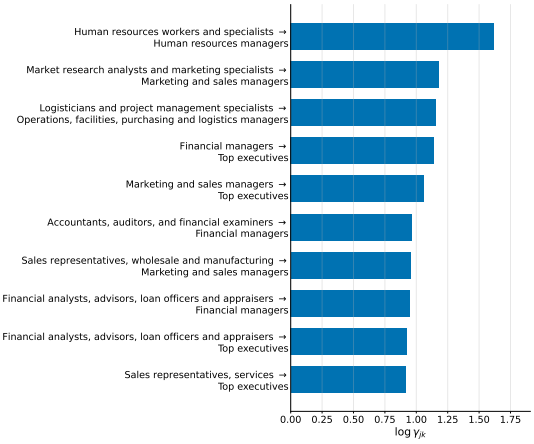
Figure A.14: Largest inferred learning gains: bachelor's-degree types



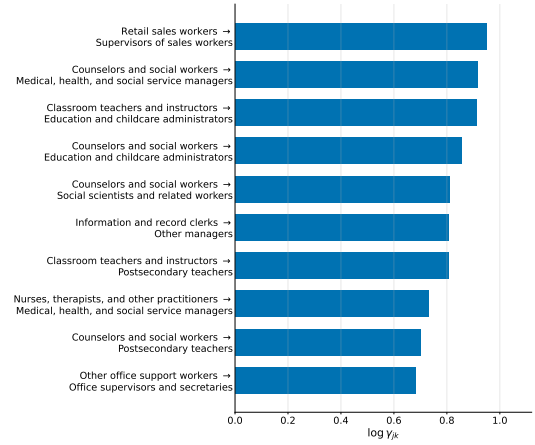
(a) Engineering, computing, and math



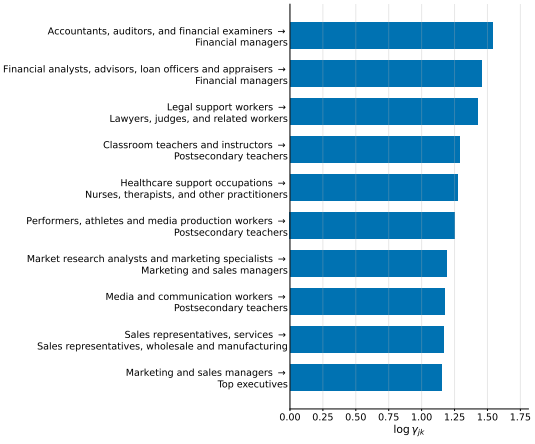
(b) Health and life sciences



(c) Business, economics, and government



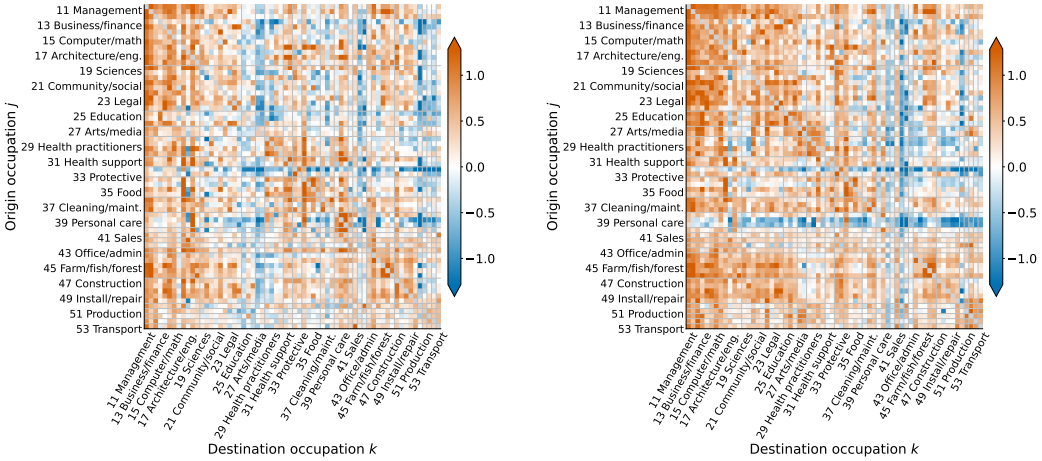
(d) Care and education



(e) Media, letters, and public service

Notes: Each panel reports the ten eligible off-diagonal occupation pairs with the largest inferred $\log \gamma_{mjk}$ for the indicated worker type. Arrows run from the initial occupation to the destination occupation. Eligible paths account for at least 0.1% of the type's young cohort, and each endpoint accounts for at least 0.5% of its young employment. Origin and destination components are not removed.

Figure A.15: Inferred learning across occupations: bachelor's-degree types



(a) Engineering, computing, and math

(b) Health and life sciences

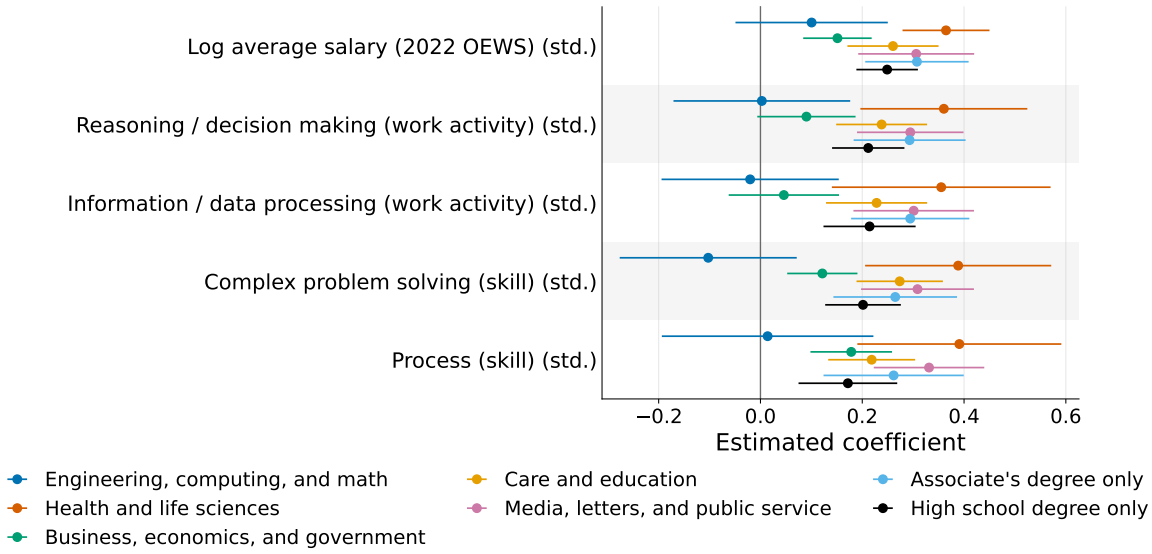
(c) Business, economics, and government

(d) Care and education

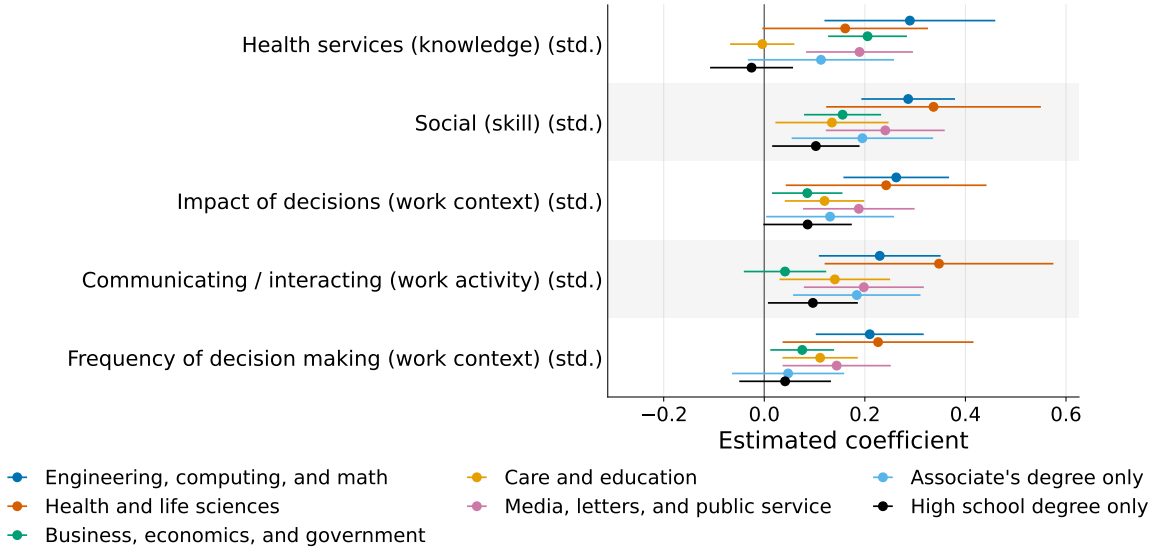
(e) Media, letters, and public service

Notes: Each panel reports inferred learning $\log \gamma_{mjk}$, with origin occupations in rows and destination occupations in columns. Occupations are grouped by SOC major category. Values are not demeaned. All panels use a common color scale, with limits set at the 98th percentile of absolute learning values across open off-diagonal cells and types; values beyond these limits receive the endpoint colors. Own-occupation learning and pairs involving occupations closed to a type are masked in gray.

Figure A.16: Directional patterns in inferred learning



(a) Measures with the largest associations (across all types)



(b) Additional measures with the largest associations for the engineering, computing, and math type

Notes: Points report coefficients from separate type-specific regressions of the directional learning contrast, $\log \gamma_{mj\ell} - \log \gamma_{m\ell j}$, on the destination-minus-origin difference in each occupational characteristic. Characteristics are standardized across occupations, and occupation pairs are weighted by their combined transition mass in both directions. A positive coefficient indicates greater inferred learning for moves towards occupations with higher values of the characteristic than for the corresponding reverse moves. Panel (a) selects measures using the pooled regressions; panel (b) selects additional measures using the engineering type's regressions, excluding those already shown in panel (a). Neither specification includes origin or destination fixed effects. Bands show 95% confidence intervals based on dyadic-robust standard errors, allowing dependence between pairs sharing an occupation. Standard errors do not account for estimation uncertainty in recovered learning.

Table A.4: Occupational similarity and inferred learning: joint regressions

	Engineering, computing and math	Health and life sciences	Business, economics and government	Care and education	Media, letters and public service	Associate's degree only	High school degree only
Knowledge distance (std.)	-0.134***	-0.106***	-0.157***	-0.055***	-0.109***	-0.152***	-0.114***
	(0.027)	(0.033)	(0.039)	(0.021)	(0.025)	(0.026)	(0.021)
Industry-mix distance (std.)	-0.104***	-0.148***	-0.098***	-0.184***	-0.131***	-0.132***	-0.094***
	(0.017)	(0.028)	(0.012)	(0.014)	(0.013)	(0.013)	(0.007)
Task-share distance (std.)	-0.048***	-0.037*	-0.064***	-0.043**	-0.057***	-0.040***	-0.041***
	(0.017)	(0.019)	(0.017)	(0.018)	(0.015)	(0.012)	(0.009)
Skill distance (std.)	-0.005	-0.020	0.023	-0.027*	0.001	-0.009	-0.017*
	(0.018)	(0.019)	(0.018)	(0.015)	(0.014)	(0.016)	(0.010)
Same SOC major group (0/1), log points	-0.057*	-0.168**	-0.105**	-0.071*	-0.065*	-0.105***	-0.101***
	(0.032)	(0.075)	(0.044)	(0.038)	(0.039)	(0.037)	(0.029)
Within R^2 (joint)	0.55	0.52	0.48	0.64	0.54	0.66	0.67
Pairs	4,032	4,032	4,032	4,032	4,032	3,660	3,660

Notes: Each column reports a separate regression for one worker type. The dependent variable is $\log \gamma_{mj\ell}$. All listed covariates enter jointly, alongside origin and destination fixed effects. Regressions use all eligible off-diagonal occupation pairs, weighted by their type-specific transition mass. Distance coefficients are reported per weighted standard deviation of the corresponding regressor after removing the fixed effects. The same-SOC-major-group coefficient is reported for a change from zero to one. The within-R^2 measures the share of variation remaining after removing the fixed effects that is explained jointly by the listed covariates. Standard errors in parentheses are two-way clustered by origin and destination occupation and do not account for estimation uncertainty in recovered learning. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

D Labor demand appendix (Section 6)

D.1 Construction of occupation-level capital shares

Following the allocation approach of Freund and Mann (2026), I construct initial occupation-level capital shares by combining the 2022 Occupational Employment and Wage Statistics (OEWS) with industry-level factor payments from the BEA–BLS Integrated Industry-Level Production Accounts. I hold the OEWS weights fixed at their 2022 values and take arithmetic averages of nominal industry-level factor payments over 2010–2019.

For each occupation j and industry i, I construct the wage bill as OEWS employment multiplied by the mean annual wage, and define s_{ji} as occupation j's share of industry i's represented private-sector wage bill. I allocate both average labor compensation, $\overline{C}_i$, and average capital income, $\overline{K}_i$, using these weights. The resulting occupational capital share is:

$$s_j^K = \frac{\sum_i s_{ji}\overline{K}_i}{\sum_i s_{ji}(\overline{C}_i + \overline{K}_i)}.$$

For grouped occupations, I sum allocated income levels before calculating this ratio; I do not average the capital shares of their constituent occupations. In the baseline, capital income comprises equipment and software. I exclude the source's combined category of structures, land, and inventories, which cannot be separated into its individual components. The denominator is therefore labor compensation plus retained capital income, rather than total national-accounts value added. Because it is reconstructed from these retained components, I apply no additional owner-occupied housing subtraction. Suppressed or unmapped OEWS cells are omitted, and allocation weights are normalized over represented occupations within each industry. The estimates consequently depend on the assumption that represented private wage bills provide suitable allocation weights for KLEMS factor payments, whose labor-compensation concept is broader than OEWS payroll.

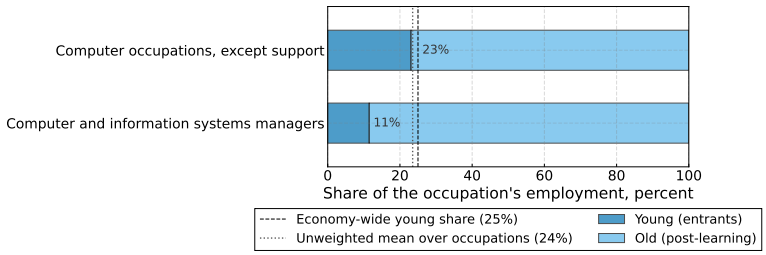
I assess sensitivity to the averaging window by repeating the construction using 2010–2022 and 2014–2019 KLEMS data while holding the 2022 OEWS weights fixed. I also compare the baseline with a total-capital specification using value added net of nonfarm owner-occupied housing, and examine the consequences of pooling real estate and rental/leasing into a single sector. Additional exercises under the broad-sector specification vary government coverage and replace apportioned KLEMS labor compensation with observed OEWS wage bills. The latter includes a 2022-only KLEMS comparison so that numerator and denominator refer to the same year. Finally, given initial occupation-level wage-bill shares ω_o, occupational expenditure shares satisfy:

$$\pi_o = \frac{\omega_o / s_o^H}{\sum_k \omega_k / s_k^H},$$

where $s_o^H = 1 - s_o^K$ is the labor share in occupation o. In ongoing work, I am improving the mapping of capital shares in the model to the data.

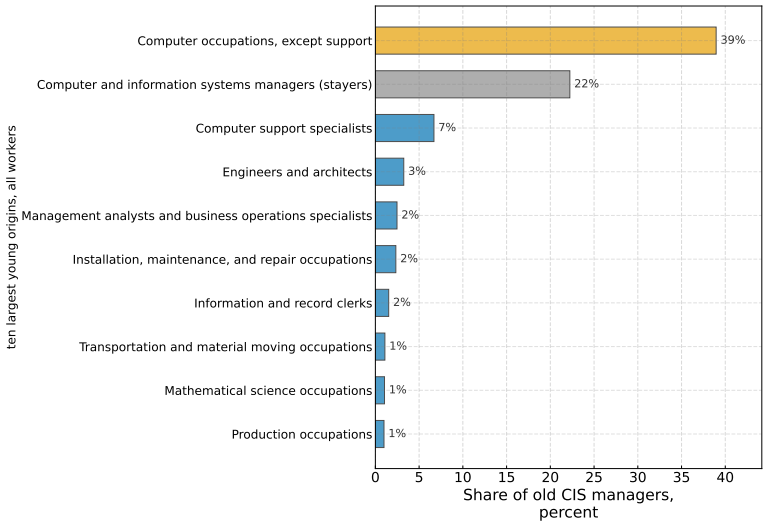
D.2 Additional figures for Section 6.2

Figure A.17: Age distribution of computer occupations and CIS managers



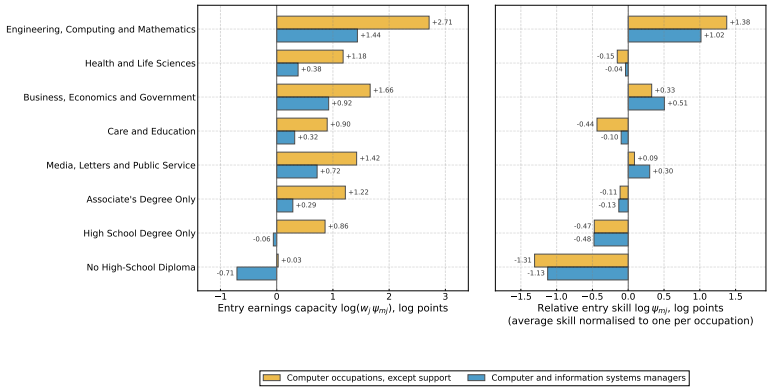
Notes: Figure shows the age distribution of computer occupations and computer and information systems managers in the initial equilibrium.

Figure A.18: Origin share of CIS managers



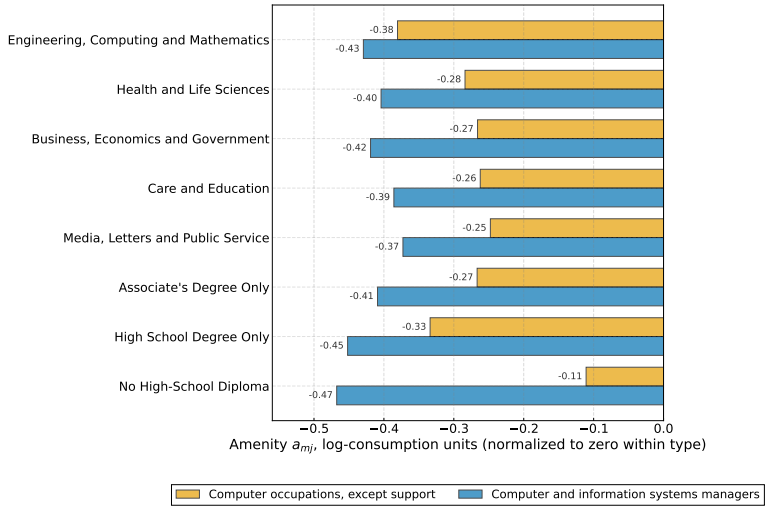
Notes: Figure shows the origin share of old workers employed as CIS managers.

Figure A.19: Entry earnings capacities



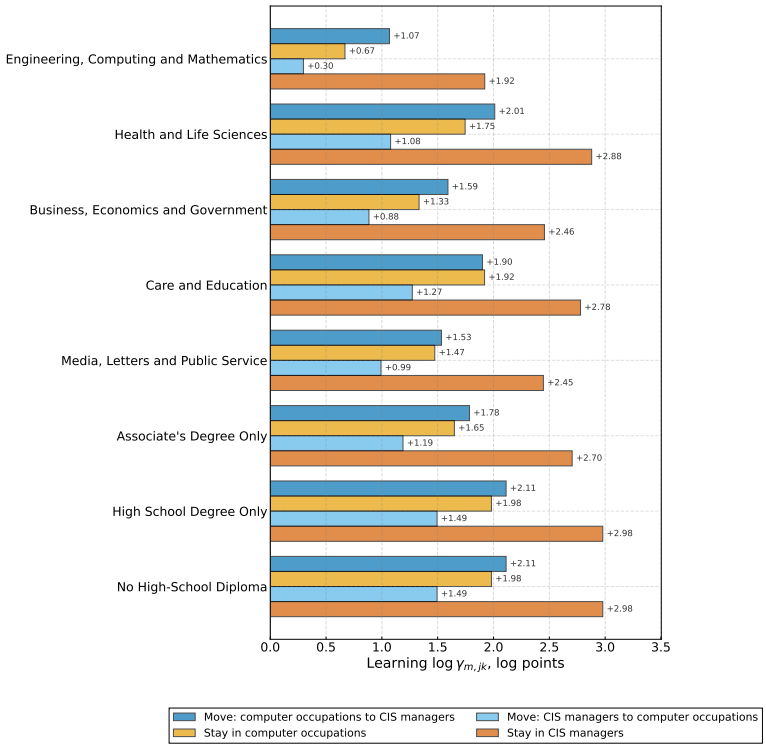
Notes: Figure shows the initial earnings capacities in computer occupations and CIS managers for all types.

Figure A.20: Amenities



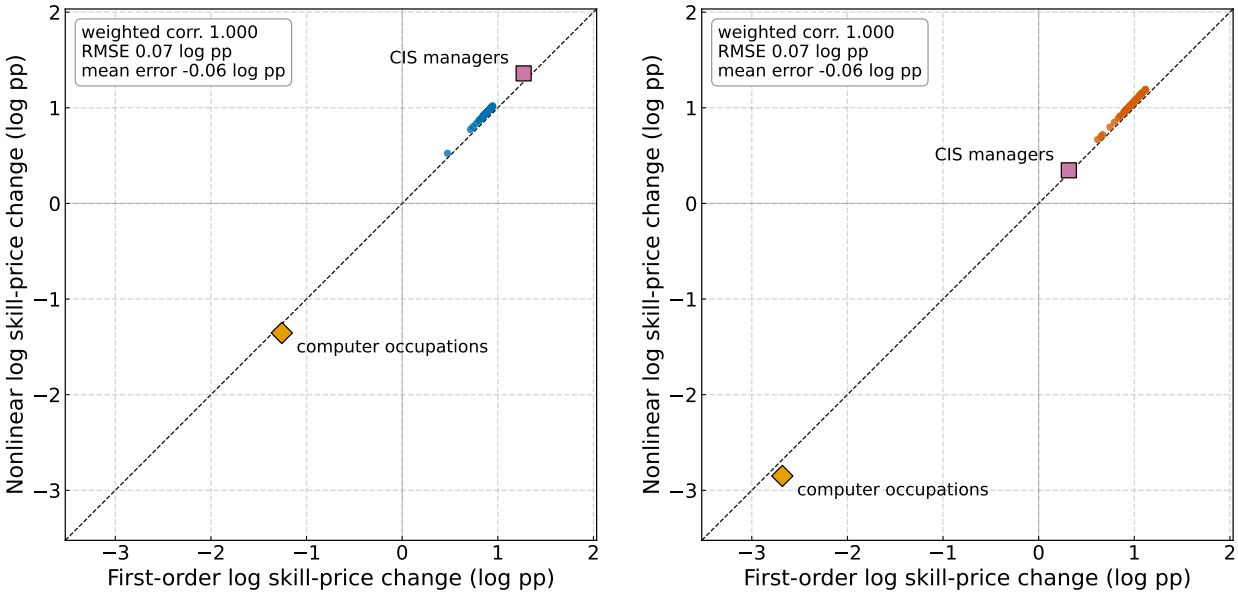
Notes: Figure shows the amenities in computer occupations and CIS managers for all types.

Figure A.21: Recovered learning



Notes: Figure shows the recovered learning technology by type.

Figure A.22: First-Order and Nonlinear Occupational Wage Responses to Computer Automation

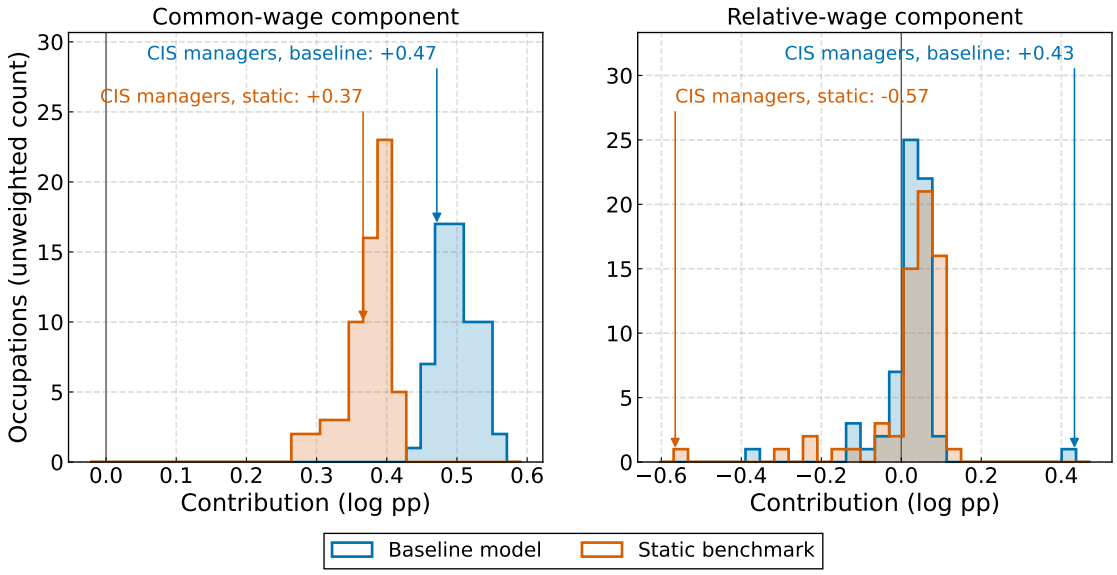


(a) Baseline model

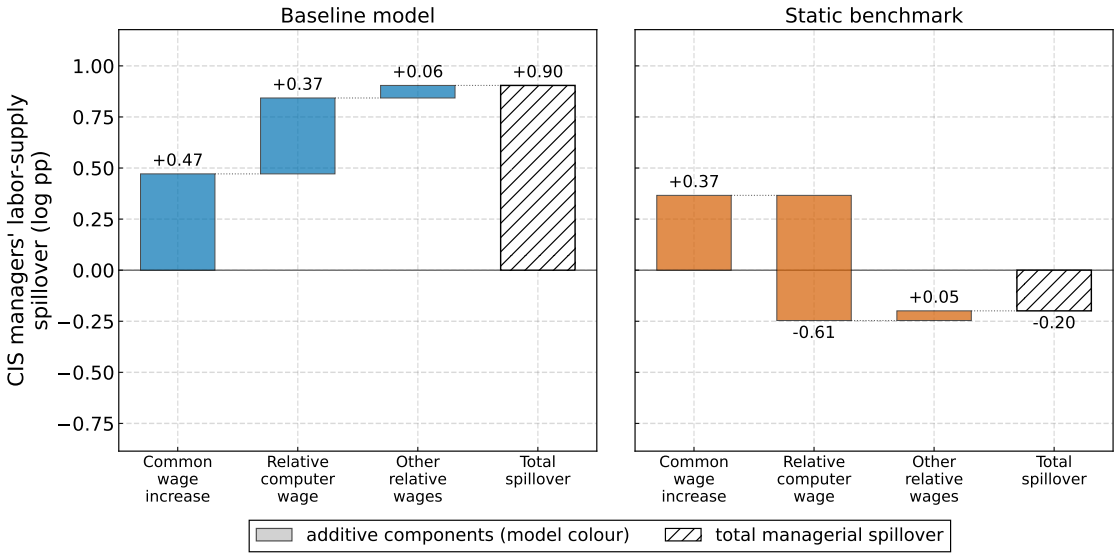
(b) Static model

Notes: Each point represents an occupation. The panels compare the first-order approximation to the fully nonlinear equilibrium change in occupational skill prices in the learning and static economies.

Figure A.23: Labor-supply spillovers from computer automation



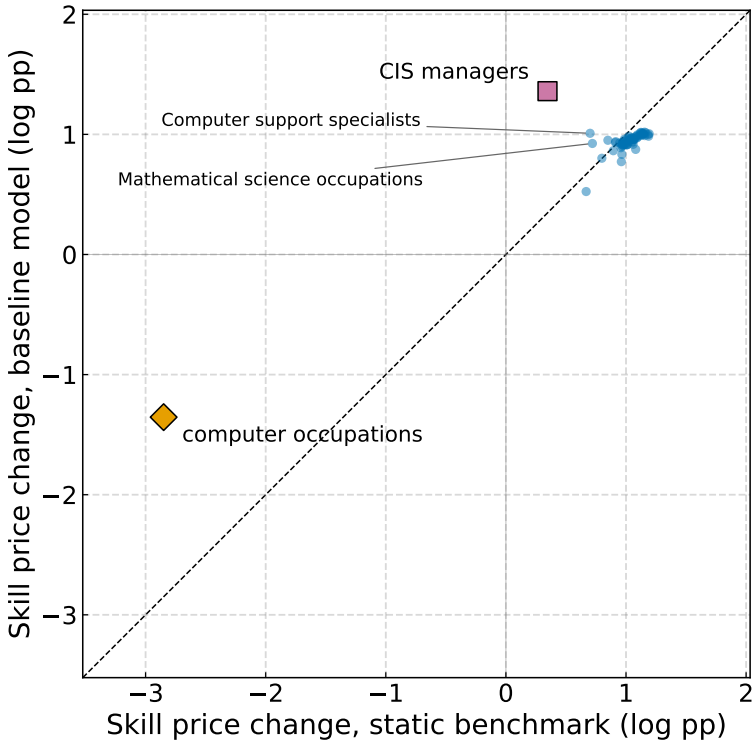
(a) Common and relative-wage components across occupations



(b) Sources of the managerial labor-supply spillover

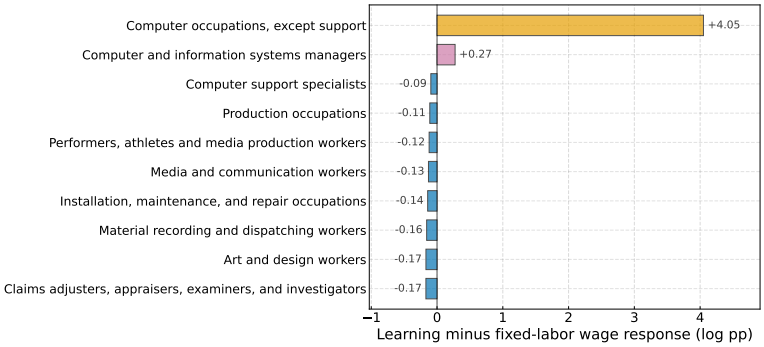
Notes: The figure decomposes the labor-supply spillover contribution to first-order occupational skill-price changes following computer automation. The common wage change is the initial-wage-bill-weighted average; relative wage changes are deviations from this average. Panel (a) shows unweighted distributions across occupations, with computer and information systems managers marked in each model. Panel (b) decomposes the managerial spillover into the common-wage contribution, the contribution from the relative computer wage change, and contributions from other occupations' relative wage changes. These three components sum to the total managerial spillover, not the full managerial skill-price response. Blue denotes the learning economy and orange the static benchmark. Contributions are measured in log percentage points.

Figure A.24: Occupational skill-price responses: learning versus static benchmark



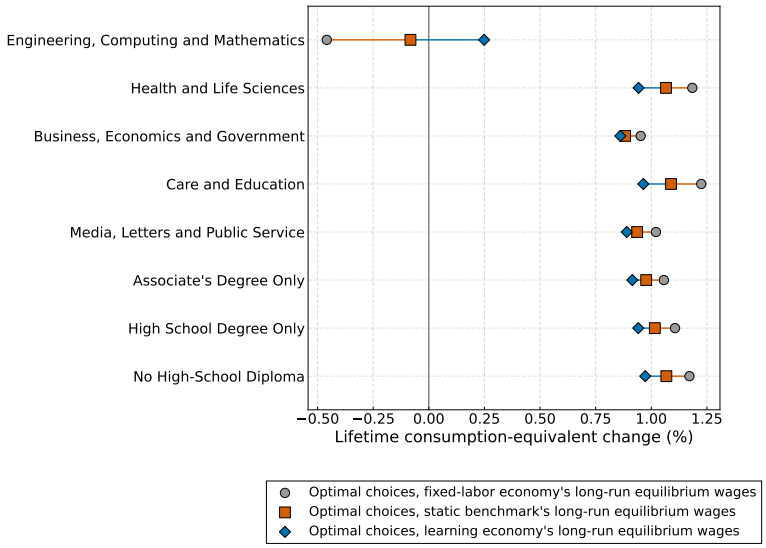
Notes: Each point represents an occupation. The horizontal axis reports the skill-price change in the static benchmark, and the vertical axis reports the corresponding change in the baseline learning economy, following the same reduction in the price of capital used in computer occupations. The dashed 45-degree line indicates identical responses across the two models. Computer occupations and computer and information systems managers are highlighted. Changes are expressed in log percentage points (100 times the change in the log skill price).

Figure A.25: Skill-price effects of labor-supply adjustment following computer automation



Notes: Bars show the difference between the long-run skill-price change in the baseline learning economy and the change when effective labor supply is held fixed in each occupation, under the same automation shock. Positive values indicate that labor-supply adjustment raises the skill price relative to the fixed-labor response, not necessarily relative to its initial level. The figure displays the occupations with the ten largest (signed) differences. Differences are expressed in log percentage points.

Figure A.26: Incidence of computer automation across workers under static benchmark wages



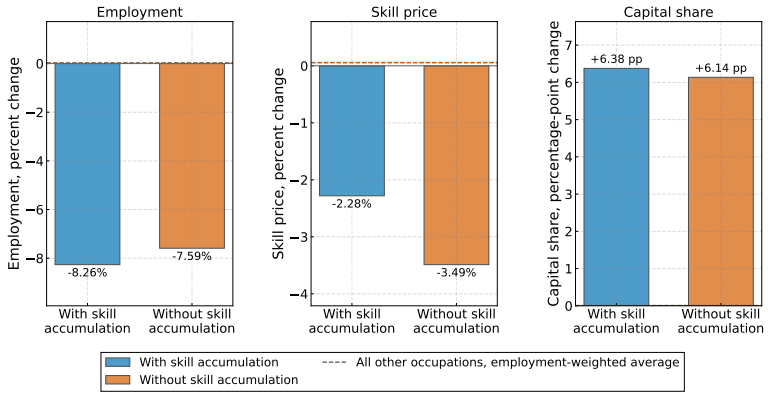
Notes: Figure reports changes in young workers' expected lifetime utility by type, expressed as consumption-equivalent percentage changes relative to the initial equilibrium. Diamonds use the learning economy's long-run equilibrium wages following the computer automation shock; squares use the static benchmark's long-run equilibrium wages under the same shock. In both evaluations, workers make optimal initial and subsequent occupational choices in the learning model, with preferences and learning technology unchanged. The comparison therefore isolates the welfare implications of the different equilibrium wage responses; it does not compare workers' welfare across different models.

D.3 Automation shocks within career ladders in Section 6.2

Figure A.27: Effects of automation in legal support occupations (exposure gap: 29.1 percentage points)



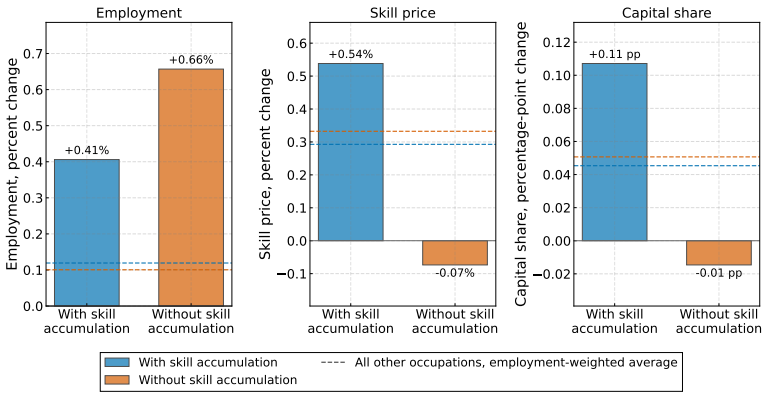
(a) Lawyers



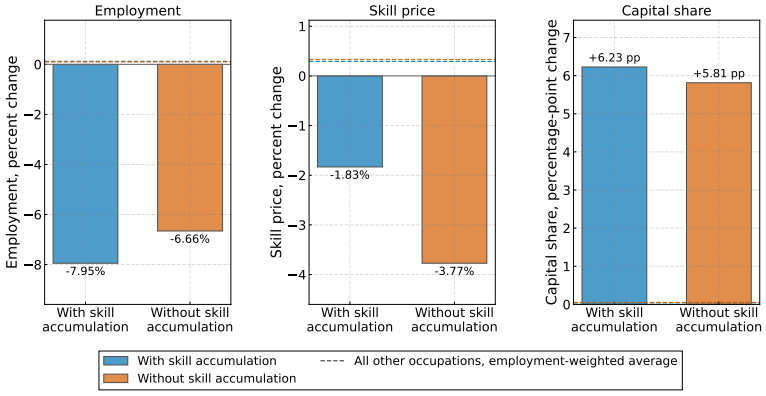
(b) Legal support workers

Notes: Figure compares the long-run responses of employment, occupational skill prices, and capital shares in the learning economy and static benchmark to a 31% reduction in the price of capital used in legal support occupations. Panel (a) reports responses for lawyers; panel (b) reports responses for legal support workers. Dashed lines show the initial-employment-weighted average response of other occupations in each economy, excluding both occupations shown. The exposure gap is exposure among legal support workers minus exposure among lawyers, measured as the percentage-point difference in the share of tasks rated as automatable.

Figure A.28: Effects of automation in accounting occupations (exposure gap: 13.4 percentage points)



(a) Financial managers



(b) Accountants, auditors, and financial examiners

Notes: Figure compares the long-run responses of employment, occupational skill prices, and capital shares in the learning economy and static benchmark to a 28% reduction in the price of capital used by accountants, auditors, and financial examiners. Panel (a) reports responses for financial managers; panel (b) reports responses for accountants, auditors, and financial examiners. Dashed lines show the initial-employment-weighted average response of other occupations in each economy, excluding both occupations shown. The exposure gap is exposure in accounting occupations minus exposure among financial managers, measured as the percentage-point difference in the share of tasks rated as automatable.

Figure A.29: Effects of automation in financial analyst occupations (exposure gap: 15.6 percentage points)



(a) Financial managers

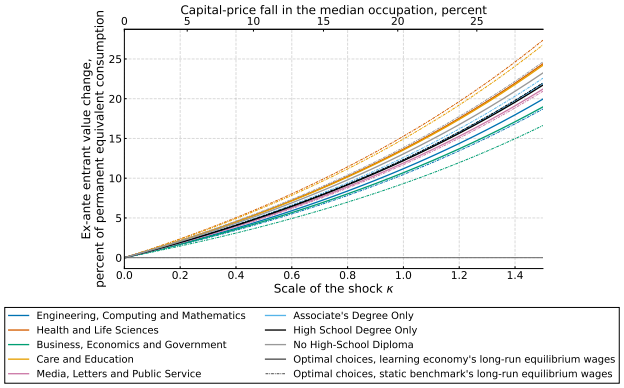


(b) Financial analysts

Notes: Figure compares the long-run responses of employment, occupational skill prices, and capital shares in the learning economy and static benchmark to a 29% reduction in the price of capital used in financial analyst occupations. Panel (a) reports responses for financial managers; panel (b) reports responses for financial analysts. Dashed lines show the initial-employment-weighted average response of other occupations in each economy, excluding both occupations shown. The exposure gap is exposure among financial analysts minus exposure among financial managers, measured as the percentage-point difference in the share of tasks rated as automatable.

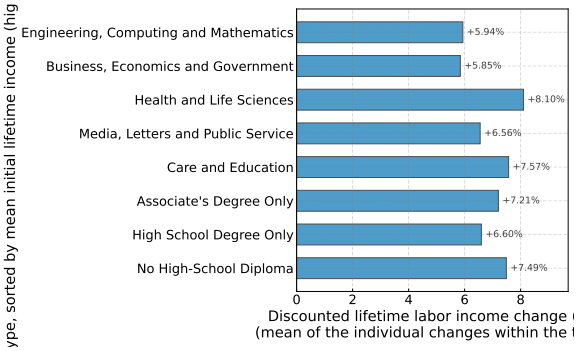
D.4 Additional figures for Section 6.3

Figure A.30: Average welfare gains for each worker type for different values of κ



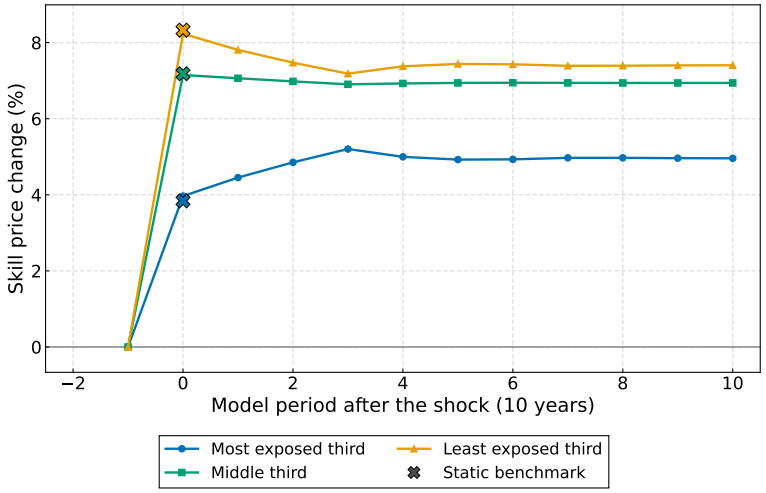
Notes: Figure reports consumption-equivalent changes in ex-ante expected lifetime utility for young workers of each type as the automation shock's scale, κ, increases. Solid lines evaluate welfare at the learning economy's long-run equilibrium skill prices; dash-dotted lines use the static benchmark's long-run equilibrium skill prices under the same automation shock. In both cases, workers choose occupations optimally according to the learning model, holding its preferences, initial skills, and learning technology fixed. Thus, differences between the curves reflect the welfare implications of the two economies' equilibrium skill-price responses, rather than differences in workers' decision problems. Welfare changes are measured relative to the initial steady state. The upper axis reports the implied percentage reduction in the capital price for the median-exposure occupation.

Figure A.31: Lifetime labor-income gains from AI automation by worker type

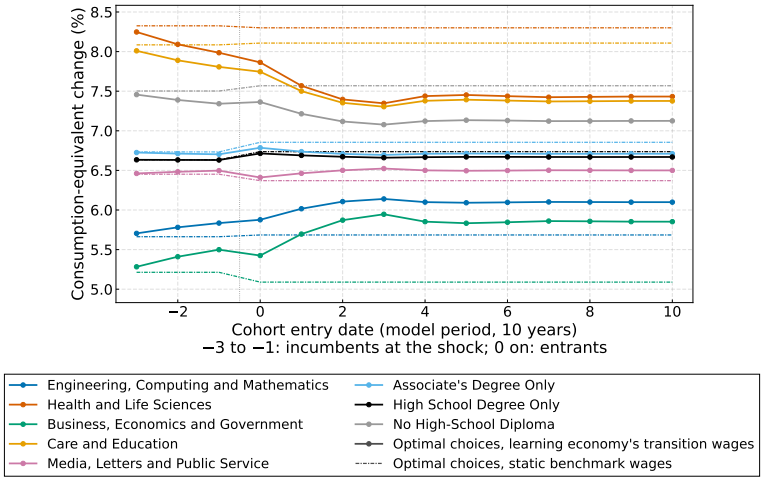


Notes: The figure reports the mean individual percentage change in discounted lifetime labor income within each worker type, comparing the initial and post-shock steady states of the learning economy. Types are ordered by mean initial lifetime income, with the highest at the top. Lifetime income sums discounted wages over four ten-year periods, holding hours fixed. Workers retain the same productivity draws across scenarios but adjust their career choices optimally. These are comparisons of entire careers under stationary equilibrium prices, not outcomes along the transition. Income excludes amenities and is distinct from consumption-equivalent welfare.

Figure A.32: Transitional dynamics following broad-based automation shock



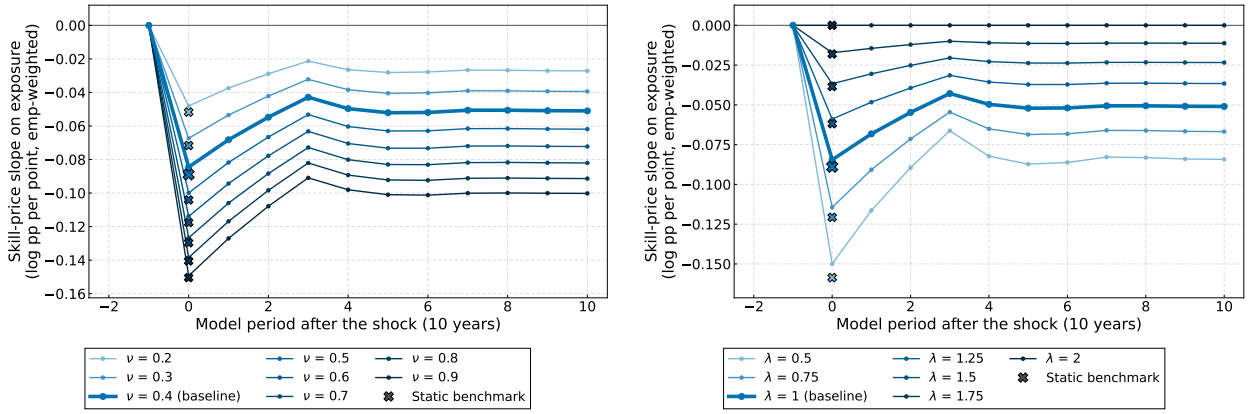
(a) Skill-price responses by exposure



(b) Welfare effects by worker type and cohort

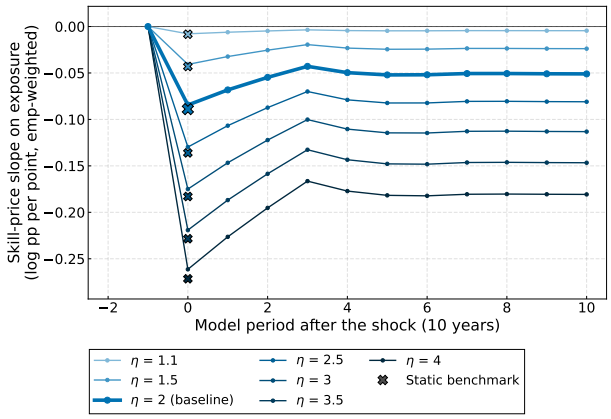
Notes: Figure shows responses to an unexpected, permanent automation shock at date zero, followed by a perfect-foresight transition. Each model period spans ten years. Panel (a) reports initial-employment-weighted averages of skill-price changes, measured as 100 times the log change from the initial steady state. Occupations are grouped into approximate thirds of initial employment by AI exposure. Solid lines show the learning economy's transition; dashed lines show the static benchmark's long-run responses. Panel (b) reports consumption-equivalent welfare changes relative to the no-shock economy. Negative entry dates identify incumbent cohorts, whose welfare is evaluated over their remaining lifetimes; nonnegative dates identify new entrants. Solid lines use the learning economy's transition wages, while dash-dotted lines evaluate the same cohorts at the static benchmark's constant post-shock equilibrium wages.

Figure A.33: Sensitivity of skill-price responses to model parameters



(a) Dispersion of productivity shocks: ν

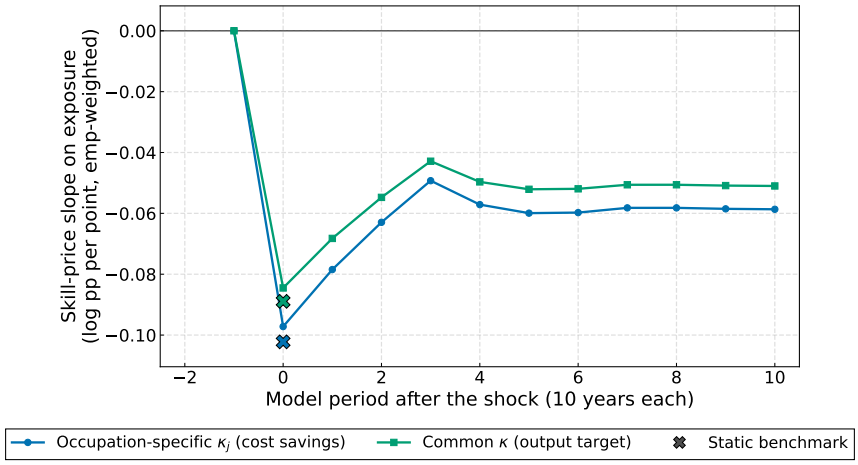
(b) Elasticity of substitution across occupations: λ



(c) Elasticity of substitution across capital and labor: η

Notes: Figure shows the sensitivity of the relationship between occupational AI exposure and skill-price responses along the transition following the automation shock. At each date, I regress changes in log skill prices on occupational exposure, weighting occupations by their initial employment shares. Slopes are measured in log percentage points per percentage point of exposure. Each panel varies one parameter, holding the other two at their baseline values: $\nu = 0.4$, $\lambda = 1$, and $\eta = 2$. The occupation-specific automation shock is held fixed across specifications. For each value of ν, the labor supply model is recalibrated to match the same initial employment and transition shares. Thicker lines indicate the baseline calibration. Crosses at period zero show the corresponding static benchmarks' slope. The shock occurs at period zero, and each model period represents ten years.

Figure A.34: Evolution of the relationship between automation exposure and skill-price responses



Notes: Figure compares the learning economy's transition under two automation shock constructions: the baseline common κ, calibrated to an aggregate output target, and occupation-specific κ_j, calibrated to cost savings on exposed tasks. At each date, I regress skill-price changes on occupational automation exposure, weighting occupations by their initial employment shares. Slopes are measured in log percentage points per percentage point of exposure. Lines with square markers show the common-κ specification; lines with circular markers show the occupation-specific-κ_j specification. Crosses at period zero show the corresponding static benchmarks' long-run slopes, with colors matching the respective shock constructions. Each shock is permanent and unanticipated at period zero. Each model period represents ten years.