Workforce Age Structure and Frontier Technology Adoption: Evidence from New Software Releases*

Tra Nguyen Sabrina Peng

Yale University

This version: October 2026

See latest version

Abstract

This paper provides causal evidence that aging slows technology adoption and sheds light on why. Using Lightcast job postings, we track local adoption of 15 software technologies following their releases, instrumenting local age composition with a Bartik-style measure based on historical age structure and national survival rates. Employers in older labor markets post increasingly fewer vacancies requiring the newest software: at year 4, a one-standard-deviation increase in the older-worker share reduces new-technology postings by about 15% of the mean pre-release legacy postings share. The gap is also larger for more difficult technologies, which we measure using new LLM-based difficulty indices. We then develop a model to separate three potential channels through which aging affects technology adoption: shorter work horizon, legacy technology experience loss, and reduced learning efficiency. Removing experience loss reduces the age adoption gap by about 88%, more than removing differences in horizons or learning efficiency. The model also shows that policies affecting workforce age structure have different implications for adoption. Relative to replacement-level fertility with retirement at 70, and assuming zero net migration, a total fertility rate of 1.6 lowers aggregate adoption by 7.5%, whereas raising the retirement age to 75 has almost no effect.

Keywords: workforce aging, technology adoption, skill demand, job postings, human capital

JEL codes: J11, O33, J24, M51

*We dedicate our most sincere thanks to Mark Rosenzweig, Orazio Attanasio, and Barbara Biasi for their invaluable guidance and support throughout this project. We also thank Joe Altonji, Costas Meghir, Michael Peters, Pascual Restrepo, and Fabrizio Zilibotti for helpful comments and conversations, and participants at the Yale Department of Economics Workshop for useful discussions. All remaining errors are our own.

1 Introduction

Technology adoption is one of the main engines of economic growth. Cross-country variation in technology adoption accounts for at least 25 percent of differences in income per capita (Comin and Hobijn, 2010). But adoption is not automatic: it usually entails substantial investment by requiring workers to acquire new skills and change their workflows. The willingness to make this investment, however, may depend on workers' age. First, older workers may have accumulated more experience with incumbent technologies, part or all of which is lost upon switching. Second, learning efficiency may fall with age as cognitive processing speed and fluid reasoning decline. Finally, older workers have shorter remaining working lives over which to recoup the investment (Ben-Porath, 1967). As workforces age across advanced economies, these differences may thus increasingly shape the pace of technology adoption. This is especially true for skill-complementary technologies, whose productivity gains scale with workers' proficiency. This paper asks whether workforce aging slows the adoption of such technologies and, if so, why.

The main challenge with answering this question is measurement. Technology adoption is difficult to observe as available data are usually coarse or span a short period of time. We overcome this challenge by relying on job postings. For a firm to adopt a new technology, it usually hires workers with relevant skills, so we interpret demand for frontier software skills as a revealed-demand measure of local technology adoption. This approach builds upon Contractor and Taska (2024) and Babina et al. (2024), who use the same dataset to identify software adoption events and measure AI adoption, respectively. We restrict our attention to IT software, which is strongly skill-complementary, for three reasons. First, IT is itself a major source of U.S. productivity growth, accounting for more than one-third of the growth in U.S. labor productivity from 2004 to 2012 (Byrne et al.), so the pace at which firms take up new tools has ramifications for the whole economy. Second, it is well covered by job postings, which name specific tools and thus allow us to date adoption at high frequency. Lastly, IT technologies can be precisely dated and have staggered releases, allowing us to exploit the variation in timing.

Using detailed online job postings data from Lightcast, we track firm demand for 15 major IT software technologies released between 2012 and 2016 across U.S. commuting zones from 2010 to 2019. For each modern technology, we also identify the relevant family of legacy technologies that it supplanted or augmented, enabling us to examine both aggregate demand for new technologies and substitution from old to new technologies in similar task environments. To address potential endogeneity resulting from persistent factors that affect both migration and adoption as well as reverse causality, we use a Bartik-style instru-

ment to predict the local age composition of college-educated workers using historical cohort structure and national survival rates, following Maestas et al. (2023). The instrument relies on variation long prior to the release of the technologies, which rules out reverse causality by construction and also assuages the persistence concern by limiting the remaining threat to factors prior to 2000 that are persistent enough to affect adoption decades later. We control for the most plausible such factors, including each commuting zone's 2000 industry composition and local skill supply.

Our estimates show that, within four years of release, employers in older labor markets post progressively fewer vacancies requiring the newest software technologies. At year 4, a one-standard-deviation increase in the share of college-educated workers aged 50–69 reduces new-technology postings by about 15 percent of the mean pre-release legacy postings share. We also find evidence of increasing substitution toward new technologies and away from legacy technologies in younger labor markets. Furthermore, this gap increases with the difficulty of the new technology, where difficulty is measured using an LLM-based index. The index construction is informed by a focus group of IT professionals and is validated against survey responses from computer scientists and software professionals.

To explain the empirical findings and evaluate policy counterfactuals, we develop and calibrate a parsimonious quantitative model linking workers' retraining decisions to firms' demand for frontier-technology skills. The model formalizes and quantifies the three mechanisms that can make retraining less attractive for older workers: a shorter horizon, experience loss, and reduced learning efficiency. In the calibrated model, removing experience loss reduces the adoption gap between younger and older workers by about 88 percent, equalizing horizons at 30 years reduces it by about 43 percent, and equalizing learning efficiency across ages reduces it by about 3 percent. The model also shows that policy levers changing workforce age structure can have different implications for technology adoption. Relative to an economy at replacement-level fertility with retirement at age 70, a total fertility rate of 1.6 lowers aggregate adoption by 7.5 percent, while a rate of 2.8 raises it by 7.4 percent, assuming zero net migration. Raising the retirement age from 70 to 75 at replacement-level fertility increases aggregate adoption by 0.45 percent. Longer careers encourage adoption, but the higher share of older workers nearly offsets this increase.

These findings contribute to research on the economic consequences of aging, technology adoption, and measurement using job postings. First, we provide causal evidence that workforce aging can inhibit technology adoption, identifying local workforce age composition as a novel determinant. This complements work on aging and automation by studying skill-complementary technologies. Second, we build a model that differentiates the channels through which this can happen and quantifies their contributions to the adoption gap, show-

ing that experience loss, rather than a shorter horizon or reduced learning efficiency, is the main driver, with implications for demographic policy. Finally, we contribute to the measurement literature by constructing an LLM-based difficulty index for technology transitions that can be used in other settings.

2 Related Literature and Contribution

This paper is most directly related to work on the economic consequences of aging. A recent but growing literature shows that demographic structure affects growth, innovation, entrepreneurship, and technology use, but the underpinning mechanisms are not well-understood.

The closest antecedents study how aging shapes technology and growth through related but distinct channels. For instance, Acemoglu and Restrepo (2022) show that aging induces greater automation. We differentiate from this paper by focusing not on skill-substituting technologies like automation but skill-complementary technologies. Similarly, while Crouzet et al. (2024) emphasizes a demand-side channel in which firms facing older consumers are less likely to adopt new technologies because those consumers are less likely to adapt to them, we focus on the supply side, asking whether the age composition of skilled workers affects firms' demand for production technologies. Other work shows that aging affects aggregate productivity, innovation, entrepreneurship, and labor-market adjustment (Jones, 2010; Liang et al., 2018; Aksoy et al., 2019; Maestas et al., 2023; Angelini, 2023; Mohnen, 2025), but it does not quantify the mechanisms through which aging can impact these outcomes.

A key micro-foundation for our hypothesis comes from Deming and Noray (2020), who show that returns to experience decelerate in fast-changing occupations as accumulated skills become obsolete. This suggests that more experienced workers face a friction in adopting new skills. Specifically, if a worker cannot transfer her skills with the legacy technology to a new technology, then an older worker who has accumulated extensive experience with the legacy technology may not adopt the new one.

The paper also contributes to the literature on technology adoption and diffusion. There has been a long tradition of work studying uneven diffusion across regions, pointing to factors such as learning costs, social transmission, and local conditions (Griliches, 1957; Mansfield, 1961; Foster and Rosenzweig, 1995). More recent work similarly stresses that adoption depends on firm capabilities, organizational constraints, uncertainty, and workforce composition (Bloom et al., 2021; Shahin et al., 2024). Closest to our mechanism, Chari and Hopenhayn (1991) model human capital as specific to a technology vintage, so that the accumulated skills of experienced workers themselves slow the diffusion of newer vintages,

but does not quantify its contribution relative to other channels. Overall, we contribute to this strand of literature by proposing another channel that generates local adoption friction.

Finally, this paper adds to the literature on using job postings as a high-frequency measure of skill demand and technology diffusion. In the past decade, job postings have emerged as an useful source of data for empirical researchers. For instance, Hershbein and Kahn (2018) use online vacancy data to track the evolution of skill demand over time, while Acemoglu et al. (2022) use this approach to study how AI exposure changes labor demand across different occupations. We build upon this literature to study a different margin, technology adoption.

3 Data and Measurement

3.1 Job postings and geography

The primary data come from Lightcast (formerly Burning Glass Technologies), a commercial provider of real-time labor market intelligence. Lightcast collects a near-universe of online job postings by scraping more than 220,000 online sources, including employer career pages, job boards, and posting aggregators. It then applies a two-step deduplication procedure that filters repeated advertisements both within sources and across sources. For each posting, the data record detailed information on the job title, employer, industry, occupation, required skills, educational requirements, experience level, location, and, when available, posted wages. These features allow us to track demand for software technologies across different labor markets over time. We aggregate the raw postings to the commuting-zone-by-year level and construct a stacked panel in which each observation corresponds to a commuting zone, a technology, and time since first stable release. This stacked design is conducive to an event study design that tracks how labor demand for new technologies evolves before and after release.

Throughout the analysis, job postings serve as a proxy for local technology adoption. When firms advertise for workers with a given modern technology, we interpret this as evidence that firms are using, piloting, or actively incorporating the technology into production and organizational workflows. In the text, we thus use "technology adoption" interchangeably with demand for labor. Of course, this proxy is imperfect, for there are many factors that can impact firms' decision to post vacancies. We therefore examine several complementary outcomes to assess whether this demand signal is consistent with technology adoption.

3.2 Technology Sample Construction

We construct a sample of software technologies that became broadly available between 2012 and 2016. This window is driven by the span of our job postings data over 2010 to 2019, to ensure that we have at least two years of observations before introduction and three after introduction of each technology, allowing us to study how usage diffuses.

We begin with the full universe of software terms identified in Lightcast, which contains roughly 8,800 technologies spanning a wide range of use cases, including HR, accounting, and IT. From this universe, we restrict our attention to major IT technologies by cross-checking Lightcast software terms against the union of technologies featured in the 2016–2024 Stack Overflow Developer Surveys,[1] a large annual survey series of software developers that reports commonly used programming languages, frameworks, databases, and development tools. We then further restrict the sample to technologies whose first stable public release occurred between 2012 and 2016. Note that we chose first stable release as the event of interest, not first public release, to ensure that the technology is production-ready.

Ultimately, our sample includes fifteen technologies in major categories including data processing systems, front-end frameworks, programming languages, DevOps tools, and machine learning libraries. For each modern technology, we also match it to its legacy technologies, which are incumbent tool or workflow released at least five years prior to the new software that a developer or firm would most naturally use for similar production tasks. Appendix Table 7 lists all major technologies included in our sample, along with their dates of introduction and descriptions of the comparable legacy technologies used in the analysis.

3.3 Local Characteristics

To characterize baseline workforce composition, we use annual American Community Survey (ACS) microdata to compute a set of demographic characteristics of interest at the local level, using information on workers' age, education, industry and occupation of employment, earnings, and public-use microdata area (PUMA) of residence. We map PUMAs to 1990 commuting zones using David Dorn's PUMA-to-commuting-zone crosswalks (Autor and Dorn, 2013; Autor et al., 2019).[2] These data allow us to characterize regional labor markets prior to the diffusion of the technologies in our sample and to construct the main demographic variables used in the analysis.

The main demographic variable is the share of college-educated workers who are ages

[1] We start in 2016 because this is the first survey year with a separately documented public-use schema.

[2] We use the 2000-vintage PUMA crosswalk for the 2000 Census and 2005–2011 ACS files, and the 2010-vintage crosswalk for later ACS files. When a PUMA is split across commuting zones, the allocation factor apportions the relevant person weights across destination commuting zones.

50–69, motivated by the fact that these technologies are all used disproportionately in occupations that are intensive in IT skills and other high skills. We take the cutoff of 50 to approximate the 75th percentile of the age distribution of IT workers.

Figure 1 shows the geographic distribution of the age structure of the college-educated workforce in 2011. Older CZs are concentrated in some areas of the Midwest and North-east as well as Florida; younger CZs are found in some metro CZs, areas of the South and the West outside of coastal areas. Figure 2 shows the corresponding cross-CZ dispersion.

3.4 Outcome measures

Our main outcome, the modern-technology share of all ads, is the share of all job postings that mention a modern technology:

$$\frac{\text{ads in } c,t \text{ mentioning modern technology } j}{\text{all job postings}_{ct}}.$$

where c indexes commuting zones, j indexes technologies, and t indexes calendar years. This outcome normalizes by local online vacancy volume.

As secondary substitution evidence, we also construct a matched-pair modern share. For each technology, this is the number of postings that mention the modern technology, including postings that mention both modern and legacy tools, divided by the number of postings that mention either the modern or the legacy technology:

$$\frac{\text{modern-only ads}_{cjt} + \text{both ads}_{cjt}}{\text{modern-only ads}_{cjt} + \text{legacy-only ads}_{cjt} + \text{both ads}_{cjt}}.$$

Since postings for a specific software are rare occurrences, we weigh this regression by the 2010–2011 legacy matched-pair posting weights so that identification is not driven by noisy cells with little baseline exposure to the legacy technology. This outcome holds the task domain fixed and thus is indicative of how much firms substitute legacy for new technologies, but it is more sensitive to technology-specific denominator movements.

We use two additional outcomes for robustness. The first is the number of ads using modern-technology relative to local labor force (ads per 100,000 workers). This outcome accounts for local labor market size when measuring demand for modern-technology, and so captures absolute changes in adoption (rather than just changes in composition). Our other robustness outcome is the share of firms in a commuting zone-year that ask for modern technology in at least one job ad. This measure captures whether modern technology use is spreading among many employers, or if just a few technology intensive employers are hiring heavily.

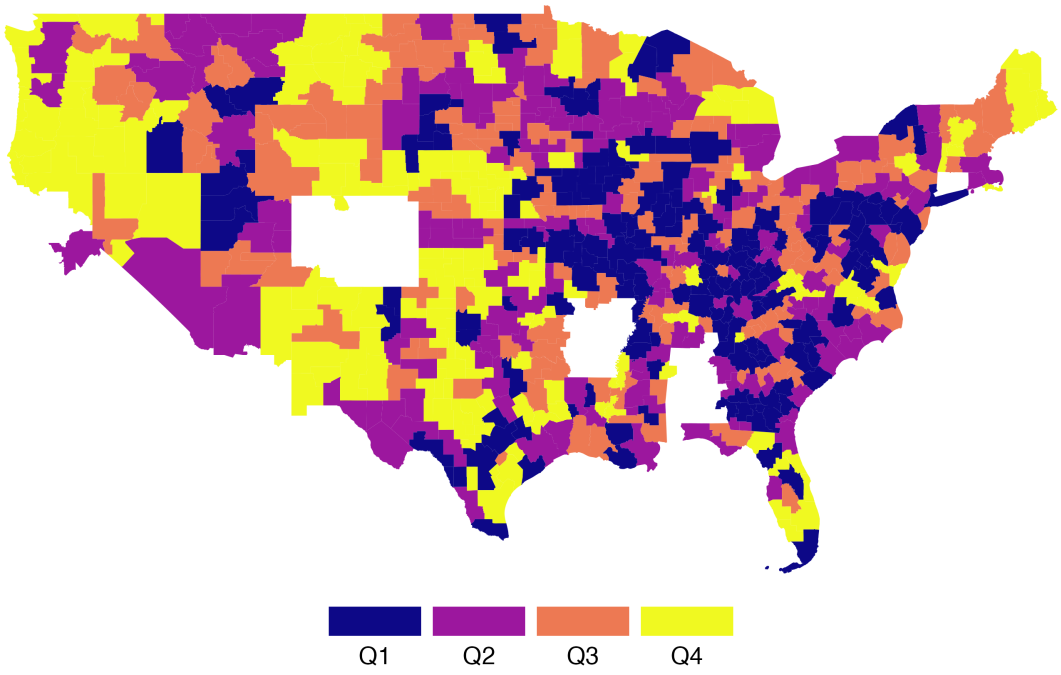


Figure 1: Geographic distribution of baseline workforce aging. Commuting zones are grouped by quartile of the 2011 share of older workers among college-educated workers.

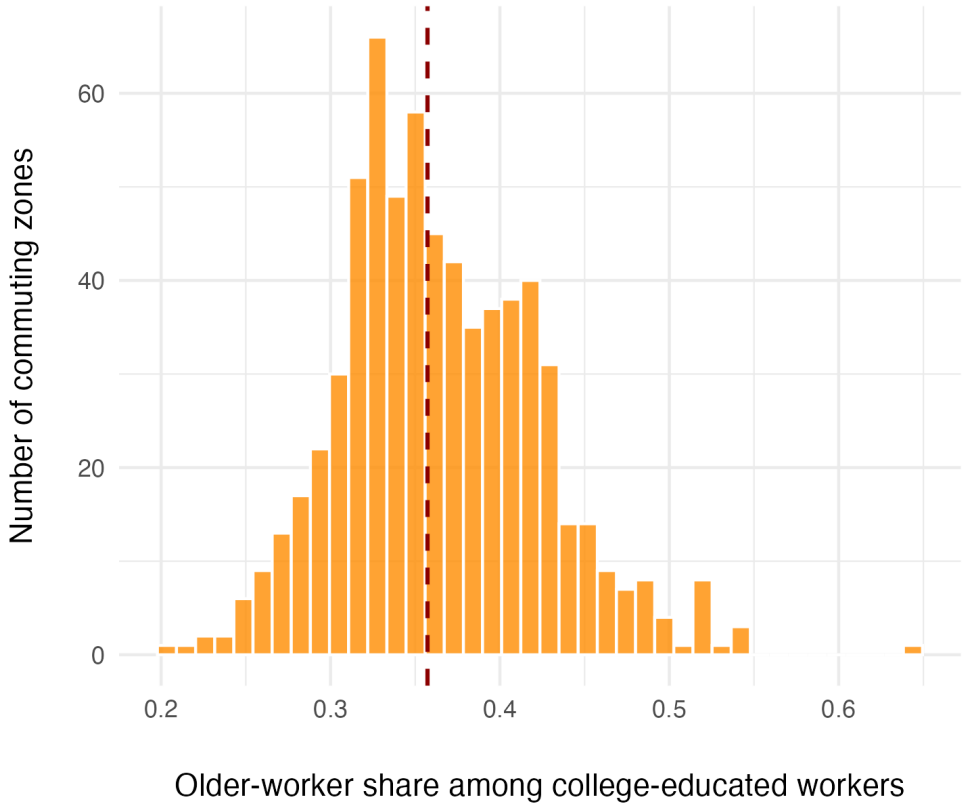


Figure 2: Cross-commuting-zone distribution of baseline workforce aging. The dashed vertical line marks the median 2011 share of older workers among college-educated workers.

To examine substitution, we match each new technology to legacy technologies used for similar production tasks and were available for production use at least five years before its release. We then measure the share of postings mentioning the new technology among postings mentioning either the new technology or its matched legacy alternatives.

4 Descriptive Evidence

As illustrated by Figure 3, postings for frontier software are negligible prior to the software release and rise sharply afterward. Nevertheless, the pace of adoption differs dramatically across commuting zones by local workforce age structure. When commuting zones are grouped by their baseline aging quartile in 2011, the youngest quartile exhibits the fastest post-release increase in modern-technology postings as a share of all postings, while the oldest quartile exhibits the slowest. The age gradient is clear in the averages across technologies.

Although this is purely descriptive evidence since older commuting zones may differ systematically from younger commuting zones in other unobserved ways relevant to technology adoption, this pattern motivates the causal analysis below.

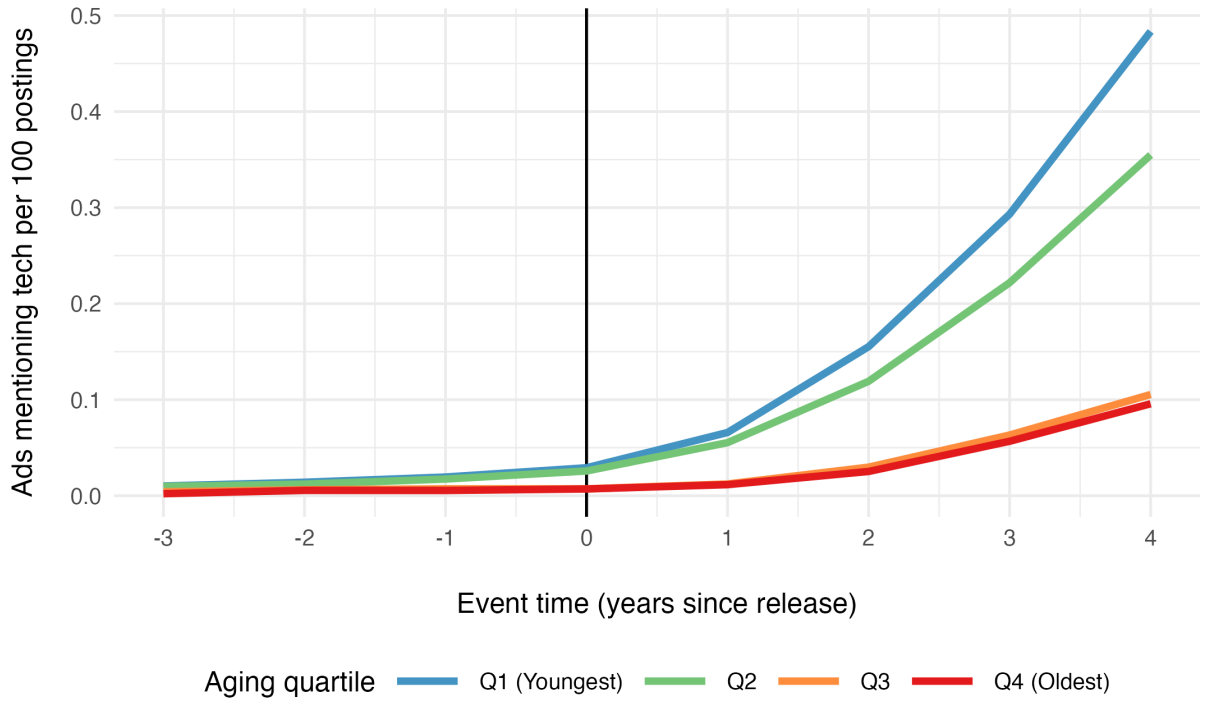


Figure 3: Modern-technology postings as a share of all postings by commuting-zone aging quartile. Lines report averages across technologies.

5 Identification Strategy

5.1 Empirical Specification

Our main empirical design is a stacked event-study that compares technology adoption around the release of each software technology across commuting zones with different baseline age structures. Let c index commuting zones, j index technologies, and t index calendar years. Define event time as

$$\tau = t - T_j,$$

where T_j is the first stable public release year of technology j. We omit event time $\tau = -1$ as the reference period. We present the event-study specification for event times $-4 \leq \tau \leq 4$. The fixed-technology-composition exercise uses its separately specified common-support window, $-3 \leq \tau \leq 3$.

Our baseline specification is

$$Y_{cjt} = \alpha_{cj} + \gamma_{j\tau} + \lambda_{s(c)t} + \sum_{\tau \in \{-4,\ldots,4\} \setminus \{-1\}} \beta_\tau \left[\mathbf{1}\{\tau_{jt} = \tau\} \times \text{OldShare}_c^{\text{col}} \right] + X'_{ct}\delta + \varepsilon_{cjt}. \tag{1}$$

The outcome Y_{cjt} is a proxy that measures local demand for frontier technology. In the main specification, it is the share of all postings that mention modern technology j. We will henceforth dub this outcome the main outcome. In additional specifications, it is either the number of postings mentioning the modern technology per 100,000 workers or a firm-level adoption margin such as the share of posting firms that mention the modern technology. Our main measure of local age composition is the share of older workers in the college-educated working-age population. We focus on the college-educated workforce because the technologies studied in this paper are disproportionately used in high-skill occupations which employ overwhelmingly college-educated workers.

We control for a rich set of fixed effects. The term α_{cj}, the CZ-by-technology fixed effect, absorbs all time-invariant differences across regions and technologies. The term $\gamma_{j\tau}$ denotes technology-by-event-time fixed effects, which absorb the common diffusion profile of each technology around release. Lastly, we control for $\lambda_{s(c)t}$, the state-by-calendar-year fixed effects, which absorb state-level shocks, policies, and business-cycle conditions that vary over calendar time. We employ state-by-calendar-year fixed effects rather than CZ-by-calendar-year fixed effects because the latter would absorb too much of the variation in the variable local age composition, as technology releases are concentrated between 2012 and 2016, resulting in substantial overlap of post-release windows in calendar time. In contrast,

state-by-year fixed effects can only remove broad state-level shocks but preserve variation across commuting zones within states. These flexible controls allow different sectors and baseline education composition to have different adoption paths over event time.

The vector X_{ct} contains 2000 industry employment shares and the 2000 college share among workers ages 20–59 in commuting zone c, each interacted with event-time indicators. The sector groups are agriculture; mining, utilities, and construction; manufacturing; wholesale, retail, and transportation; finance and real estate; professional services; information and communications; education, health, and arts; and other services and public administration. In robustness checks, we further control for university capacity and manager age structure, each interacted with event time.

Our coefficient of interest is β_τ, which measures the differential change in technology adoption at event time τ associated with a one-unit increase in the baseline old-worker share among the college-educated workforce, relative to the omitted pre-period $\tau = -1$. A negative coefficient means that commuting zones composed of older college-educated workers experience slower post-release adoption of frontier technologies. The sequence of coefficients $\{\beta_\tau\}$ traces the evolution of this adoption gap before and after release.

5.2 Instrumental Variable

The age structure was measured in 2011, before the stable release of our technologies. This should eliminate most of the concern about reverse causality, although some technologies had their initial releases a couple of years prior. A second concern remains, however. The age structure in 2011 is itself the product of migration and economic processes over the preceding decade. If the factors driving those processes also encourage technology adoption—stronger technology ecosystems, faster-growing software sectors, more research university pipelines, or more innovative firms—then they generate a spurious negative relationship between workforce age and technology adoption even in the absence of a causal effect of age structure itself. Note that this concern is distinct from firms relocating in response to the local supply of young workers, which is considered part of the equilibrium response being measured.[3] To address it, this study follows Maestas et al. (2023) and Mohnen (2025) by employing a Bartik-style demographic instrument constructed with historical age composition. Because it is built from variation predating the releases by more than a decade, the instrument completely rules out any potential reverse causality from initial releases. The only remaining threat is the possibility that the 2000 cohort structure is itself correlated with local characteristics persistent enough to still predict adoption more than a decade later, which we

[3]Hopenhayn et al. (2022) provides evidence that population growth induces firm entry.

address with the controls described below.

The intuition behind the instrument is simple. We build it by aging local college-educated cohorts by ten years using national survival rates, while keeping migration patterns fixed. Let $N^{\text{col}}_{c,h,2000}$ denote the number of college-educated individuals in commuting zone c, age bin h, in 2000, and let $S_{10}(h)$ denote the national ten-year survival rate for cohort h. For age groups whose education is already observed in 2000, the predicted size of cohort $h+10$ in 2010 is

$$\widetilde{N}^{\text{col}}_{c,h+10,2010} = S_{10}(h)\, N^{\text{col}}_{c,h,2000}.$$

For illustration, suppose the predicted number of college-educated individuals aged 50–54 in commuting zone c in 2010 is

$$\widetilde{N}^{\text{col}}_{c,50\text{–}54,2010} = S_{10}(40\text{–}44)\, N^{\text{col}}_{c,40\text{–}44,2000}.$$

One problem with the formulation above is that final educational attainment is not yet fully observed in 2000 for younger entrant cohorts. For these cases, college attainment is imputed using 2000 commuting-zone-specific college attainment for the 25–29 age group. Specifically, let p^{col}_c denote the share of individuals aged 25–29 in commuting zone c with a college degree in the 2000 Census. This rate is measured prior to the technology releases in the sample and is therefore predetermined with respect to our softwares' diffusion. For entrant cohorts, the predicted future size is :

$$\widetilde{N}^{\text{col}}_{c,h+10,2010} = S_{10}(h)\, N^{\text{all}}_{c,h,2000}\, p^{\text{col}}_c.$$

Define the college-educated working-age population as

$$\mathcal{W} = \{20\text{–}24, 25\text{–}29, \ldots, 65\text{–}69\},$$

and define older college-educated workers as

$$\mathcal{O} = \{50\text{–}54, 55\text{–}59, 60\text{–}64, 65\text{–}69\}.$$

The predicted number of older college-educated workers in commuting zone c in 2010 is

$$\widetilde{N}^{\text{col}}_{c,\mathcal{O},2010} = \sum_{h\in\mathcal{O}} \widetilde{N}^{\text{col}}_{c,h,2010},$$

and the predicted number of college-educated working-age individuals is

$$\widetilde{N}^{\text{col}}_{c,\mathcal{W},2010} = \sum_{h \in \mathcal{W}} \widetilde{N}^{\text{col}}_{c,h,2010}.$$

The instrument is defined as the predicted old-worker share among the college-educated workforce:

$$\widehat{\text{OldShare}}^{\text{col}}_{c,2010} = \frac{\widetilde{N}^{\text{col}}_{c,\mathcal{O},2010}}{\widetilde{N}^{\text{col}}_{c,\mathcal{W},2010}}.$$

Note the one-year difference between the endogenous 2011 baseline age share and the year of cohort size that the IV predicts. This discrepancy reflects a data limitation: while 2011 immediately precedes the technology releases and is thus the most appropriate baseline, we only have 10-year survival rates, not 11-year. Nonetheless, local age structure is highly persistent at the annual frequency, so this incongruence does not affect the strength of the first stage. The first stage should therefore be interpreted as relating the predicted 2010 age structure to the actual baseline age structure in 2011.

As discussed earlier, the 2000 cohort structure, or the baseline college-attainment rate used to impute entrants, may be correlated with persistent time-varying local characteristics that also predict later technology adoption. For instance, a CZ may comprise strong IT-intensive industries, or that the existing labor supply is already high-skilled. To address this concern, we condition on both a rich set of fixed effects and a rich set of covariates: each commuting zone's 2000 employment shares across broad industries and the 2000 college share among workers ages 20–59, both interacted with event times.

To reassure ourselves further, we run a balance test of our IV against a range of 2000 characteristics that potentially affect technology adoption. Table 2 reports balance diagnostics for the predicted old-college-worker share. After these controls, the instrument is not strongly related to median wage-income levels, wage-income growth, unemployment, or baseline science, engineering, and computer employment. University capacity remains significantly related to the instrument, but the magnitude is modest, so robustness checks are also reported that add university capacity as control. Although a balance check is by no means a definite test for exogeneity, it serves as a reassuring first pass.

Appendix Figure 22 presents a Rotemberg-style decomposition of this cohort-survival IV, following Goldsmith-Pinkham et al. (2020), to better understand the source of the variation in our IV. As observed, older bins contribute directly to the future old-worker share, while younger bins primarily influence the projected college-educated workforce denominator.

For each event time τ, the endogenous regressor is the interaction between the event-time

indicator and the actual old-worker share,

$$Z_{cjt\tau} = \mathbf{1}\{\tau_{jt} = \tau\} \times \text{OldShare}_c^{\text{col}}.$$

The excluded instrument is the corresponding interaction with the predicted old-worker share,

$$\widehat{Z}_{cjt\tau} = \mathbf{1}\{\tau_{jt} = \tau\} \times \widehat{\text{OldShare}}_c^{\text{col}},$$

where $\widehat{\text{OldShare}}_c^{\text{col}}$ is constructed from lagged cohort structure and national survival rates. The stacked first stage projects the full vector of endogenous event-time interactions on the full vector of excluded predicted-aging interactions:

$$\begin{aligned} Z_{cjt\tau} = & \sum_{\tau' \in \{-4,\ldots,4\} \setminus \{-1\}} \pi_{\tau\tau'} \widehat{Z}_{cjt\tau'} \\ & + X_{ct}'\rho + \alpha_{cj} + \gamma_{j\tau} + \lambda_{s(c)t} + u_{cjt\tau}. \end{aligned} \tag{2}$$

The second stage is given by equation (1), with the set of $Z_{cjt\tau}$ terms instrumented by the corresponding set of $\widehat{Z}_{cjt\tau}$ terms. Standard errors are clustered at the commuting-zone level. Although the second-stage equation is reminiscent of a staggered difference-in-differences regression, the resemblance is merely because we employ an event-study design to trace the adoption path around each technology's release. The identifying variation comes from our instrument, conditional on controls and fixed effects, not from a parallel-trends comparison across older and younger commuting zones.

Because equation (2) interacts both the endogenous age-structure variable and the excluded instrument with event time, the first stage is not a single equation but a system: one equation for each endogenous interaction $Z_{cjt\tau}$. This creates a well-known difficulty for assessing instrument strength, which Buchheim and Watzinger (2023) confront in an identical setting, instrumenting a time-invariant cross-sectional regressor interacted with date dummies. By construction, $\widehat{Z}_{cjt\tau}$ is the only excluded instrument with explanatory power for the equation whose dependent variable is $Z_{cjt\tau}$; every other event-time interaction is uninformative for that equation. Adding event times therefore mechanically depresses any statistic for the joint significance of the full instrument vector, so a single joint first-stage F understates instrument strength by an amount that grows with the number of event times. Critical values for weak-instrument tests with many instruments and many endogenous regressors are in any case unavailable (Stock and Yogo, 2005).

We therefore follow Buchheim and Watzinger (2023) and assess relevance from the cross-section. Because the instrument, the endogenous variable, and the controls are all time-

invariant objects interacted with event-time indicators, the system in equation (2) closely resembles a repeated cross-section, with one equation per event time. Its identifying variation is almost entirely cross-sectional, so the strength of the instrument is summarized by the cross-sectional relationship between the predicted and actual old-worker shares.

Table 1 reports that relationship. Alongside the F-statistic we report the Shea partial R^2 of the excluded instrument, which is far less sensitive than the F-statistic to the uninformative instruments that the event-time interactions introduce and is therefore the more stable bridge between the cross-sectional diagnostic and the full system.

Absorbing state fixed effects (Column (2)) attenuates the coefficient from 0.412 to 0.276 and reduces the Shea partial R^2 from 0.165 to 0.076. The within-state first stage, however, remains strong, with an F-statistic of 50.3, and the within-state Shea partial R^2 of 0.076. Our instrument strength is thus comparable to Buchheim and Watzinger (2023)' instrument in their most demanding specification, where they judge the instrument sufficiently strong on the same criterion.

Appendix Table 10 supplements this with equation-specific effective F-statistics for each event time, which reports instrument strength separately for each first-stage equation, using only the variation relevant to that equation.

Table 1: First-stage relationship between predicted 2010 and actual 2011 baseline age structure

	Actual old-worker share in 2011 (%)	
	(1)	(2)
Predicted old-worker share in 2010 (%)	0.412	0.276
	(0.036)	(0.039)
State fixed effects	No	Yes
R^2	0.287	0.466
Shea partial R^2	0.165	0.076
First-stage F-statistic	128.9	50.3
Commuting zones	665	663

Notes: Both the actual and predicted old-worker shares are expressed as percentages (0–100) of the college-educated workforce, so the coefficient gives the percentage-point increase in the actual 2011 share associated with a 1-percentage-point increase in the instrument. Standard errors are in parentheses. Both columns control for the baseline 2000 industry shares and the 2000 share of workers ages 20–59 with a college degree. Column (2) additionally includes state fixed effects. The two commuting zones dropped in column (2) are singletons within the state, thus removed when state FEs are included.

5.3 Results

Figures 4–8 present the preferred and secondary IV results. As observed, prior to the stable release of the technologies, the coefficients are close to zero and statistically imprecise, which is expected as the software had not been released for production use yet. However, after release, there is a divergence between the young and old commuting zones where the old commuting zones progressively lag behind, as reflected in the increasingly negative coefficients. This gap is evident across all outcomes examined. The weighted matched-pair outcome provides secondary evidence suggesting substitution away from legacy and into newer technologies in younger commuting zones.

Figure 4 plots our main outcome: modern-technology postings as a share of all job postings in the commuting-zone year. A one-standard-deviation increase in the older-worker share reduces new-technology postings at year 4 by about 15 percent of pre-release legacy postings.

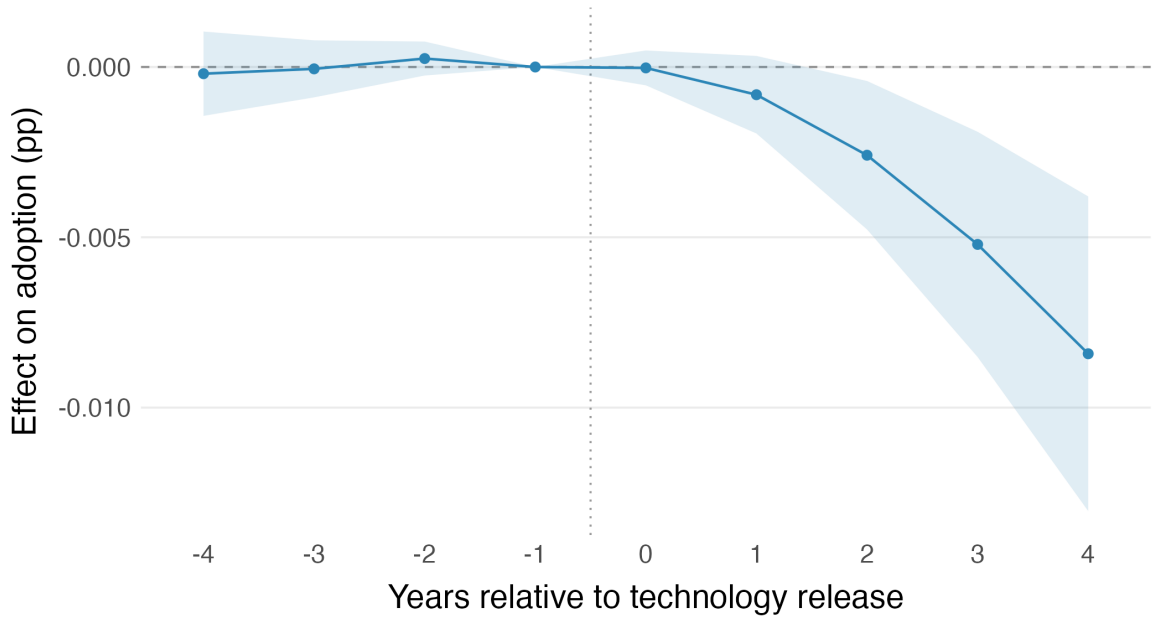


Figure 4: Event-study estimates for modern technology postings as a share of all job postings

As a robustness check, we also examine other outcomes as proxies for adoption. Reassuringly, workforce-normalized vacancy counts and the share of firms posting vacancies requiring modern technologies also show the same pattern as the main outcome. At year 4, a one-standard-deviation increase in the older-worker share implies about 4.9 fewer modern-technology postings per 100,000 workers. This is approximately 28 percent of the mean pre-release legacy-only posting rate of 17.5 per 100,000 workers. For the firm margin, the year-4 effect is a decline of 0.098 percentage points, approximately 16 percent of the mean

pre-release legacy-only posting-firm share of 0.625 percent. Both benchmarks are unweighted means across commuting-zone–technology pairs in the year before release ($\tau = -1$); the effects use the cross-CZ standard deviation of the baseline older-worker share, 5.90 percentage points. [4]

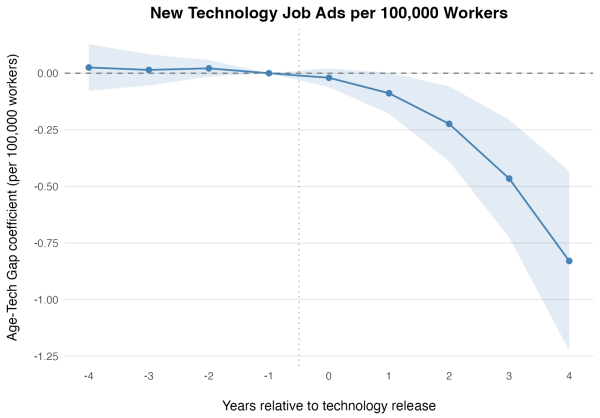


Figure 5: Modern technology postings per 100,000 workers

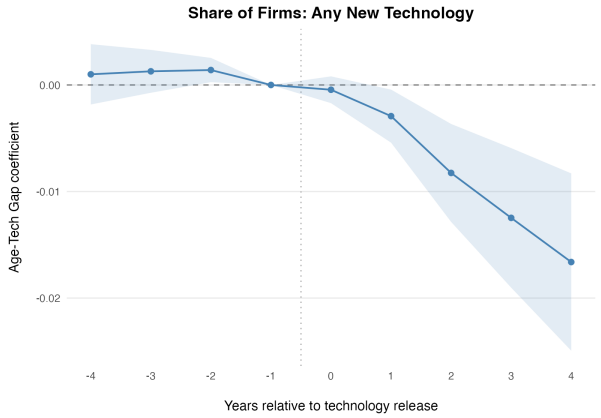


Figure 6: Share of firms demanding modern technologies

Figure 7 reports the matched-pair modern share, defined as the share of postings mentioning the modern tool among postings mentioning either the modern or legacy tool in the same technology family. This outcome compares demand for a modern tool to demand for its legacy alternative. As such, the matched-pair modern share is in some sense the ideal proxy outcome to study, although its smaller denominator means that it is more noisy in practice: just one or two job postings can greatly affect this share. Therefore, all regressions using the matched-pair outcomes are weighted by each CZ-by-technology's initial (i.e. fixed at its 2010–2011) level of legacy matched-pair job postings, which downweights CZ-by-technologies with small numbers of matched job postings (where the outcome mechanically fluctuates greatly) and puts relatively more weight on the CZ-by-technology-level labor market shares in each technology where there was meaningful exposure to the legacy tool pre-release.

Estimates based on this noisier outcome confirm the negative age gradient post-release, and thus the slower pace of substitution away from legacy towards modern tools, in older labor markets. We view evidence using the matched-pair modern share as providing evidence about substitution while we focus on the legacy, firm, posting-count, and all-ad shares to understand broader patterns of diffusion.

[4] Appendix Figure 17 reports weighted versions of the all-ad-share, posting-count, and firm-level outcomes.

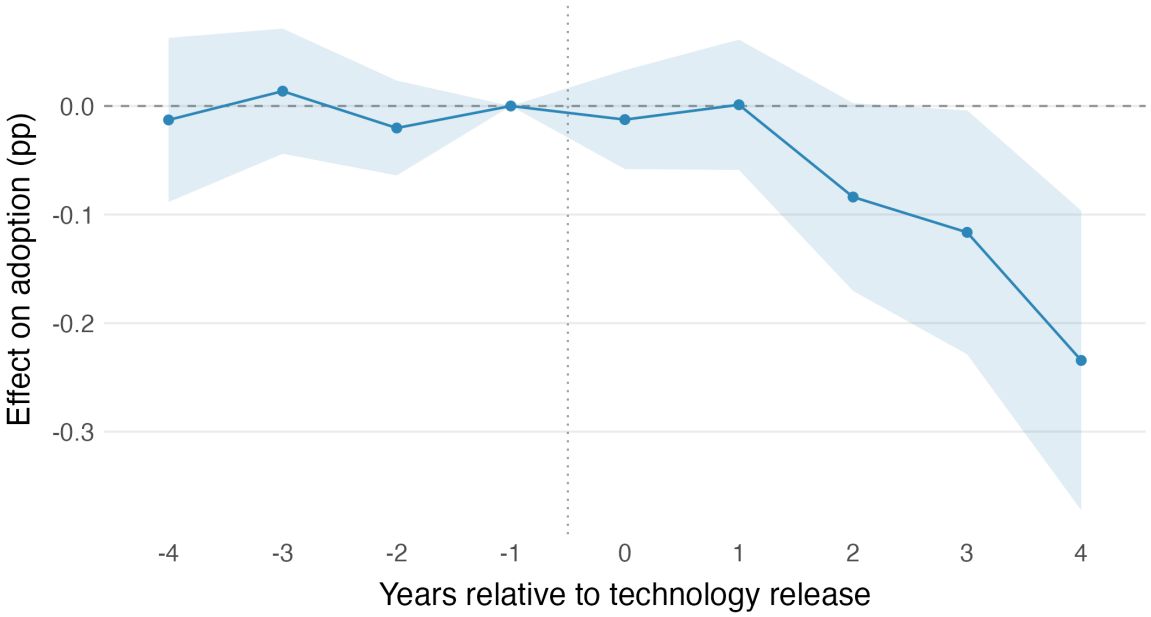


Figure 7: Weighted matched-pair IV event-study estimates with state-by-year fixed effects

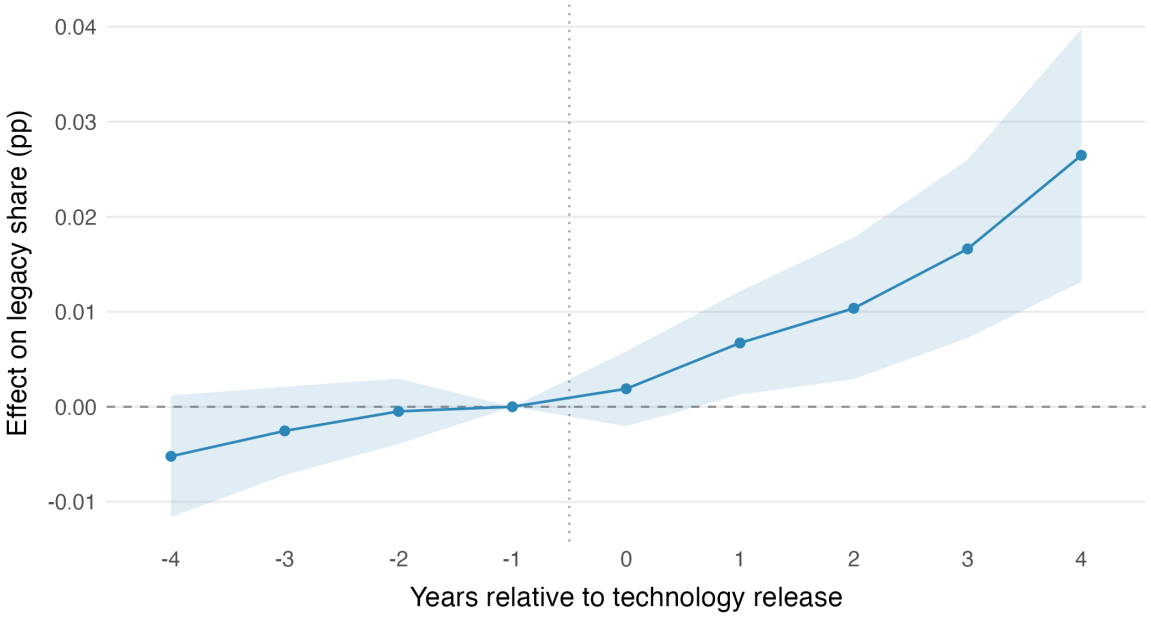


Figure 8: IV event-study estimates with state-by-year fixed effects: legacy-only share of job postings

The legacy results reinforce this interpretation. Legacy technologies are defined as technologies that were used for similar tasks as the new technologies, and predate the modern alternative by at least five years so that they are already well established by the time the new technology appears. The latter criterion ensures that our comparison is between old and new technologies, and not contemporaneous rivals. Figure 8 suggests that older commuting zones remain more reliant on these legacy technologies after the release of the modern alternative.

One concern in interpreting our empirical results is that job postings measure the flow of new hiring rather than the stock of technology already in use; therefore, older regions might appear to have weaker adoption simply because they adopted these technologies earlier and therefore need fewer additional hires once the release window begins, or because their workforce is already equipped with these new software skills and thus no longer needs to hire new workers. We argue that this is implausible: the technologies in the sample are new enough that large installed stocks before stable release are unlikely, although prerelease experimentation, piloting, or organizational preparation could still occur.

5.4 Balance and Robustness Checks

Table 2 reports selected balance diagnostics for the cohort-survival instrument. Each row reports how a one-standard-deviation increase in the predicted old-worker share is related to the listed predetermined CZ characteristic, scaled by the standard deviation of that charac-

Table 2: Selected balance diagnostics for the predicted old-worker share

Outcome	SD-scaled association	p-value	N
Log median wage income, 2000 Census	-0.020	0.451	663
Log median wage-income growth, 2000–2010	-0.027	0.451	663
Unemployment rate (ages 18+), 2000	0.014	0.712	663
Science, engineering, and computer occupation share, 2000	-0.031	0.181	663
Number of colleges and universities, 2000	-0.086	0.003	663

Notes: Each row reports a separate commuting-zone-level regression of the listed variable on the predicted old-worker share among college-educated workers. The reported association is the effect of a one-standard-deviation increase in the predicted old-worker share divided by the standard deviation of the outcome. All regressions include state fixed effects, baseline 2000 industry-share controls, and the 2000 college share among workers ages 20–59. Wage income is annual wage-and-salary income: the 2000 Census reports income for 1999, while the 2010 ACS reports income over the preceding 12 months. Growth is the difference in log median nominal wage income between the two surveys. Standard errors are heteroskedasticity robust.

teristic. The instrument is not strongly related to median wage-income levels, wage-income growth, unemployment, or baseline science, engineering, and computer employment after our controls. On the other hand, there appears to be a statistically significant correlation between our IV and university capacity, although the imbalance is rather modest: a one-standard-deviation increase in the IV is associated with about 0.7 fewer institutions.

We report two additional robustness checks targeted at the remaining balance and interpretation concerns. First, given that the university network variable is significant in the balance check, we add the number of local colleges and universities in 2000, interacted with event time, to absorb baseline university capacity and related research infrastructure. Second, we also add the 2011 share of managers age 50 and older, interacted with event time, to address the possibility that the estimated workforce-aging effect is actually that of older decision makers. Appendix Figures 19 and 21 report these checks. The all-ad-share, posting-count, and firm-level outcomes remain negative under the university-capacity robustness. In the manager-age robustness check, which uses the same set of controls as the main specification, the manager-age coefficients are close to zero and statistically insignificant, while the workforce-aging coefficients remain negative and statistically significant for these outcomes.

6 Mechanism

What explains the gap in diffusion between old and young commuting zones? The most obvious explanation, going back to Ben-Porath (1967), is that older workers have shorter

remaining working lives over which to recoup potential adoption benefits relative to young workers. Indeed there is evidence suggestive of this mechanism. Previous literature has shown that a change in the pension system which extends workers' working life induces more on-the-job training (Montizaan et al., 2010; Fürstenau et al., 2023). Another reason is loss of experience with legacy technologies. Frontier software often requires workers to learn new conceptual abstractions, new tooling, and new workflows. Older workers, having accumulated more experience with legacy technologies, may face greater losses of accumulated legacy experience. Lastly, older workers may also face higher direct learning costs as they have reduced learning efficiency due to cognitive decline.

All three channels imply a larger gap for more difficult technologies. A more difficult technologies usually means longer transition, which leaves even less time to recoup the investment, makes the loss of a larger stock of legacy experience even longer to rebuild, and imposes a heavier burden on workers who learn more slowly. To test the presence of any of these channels, we need to construct two technology-level indices for each legacy–modern technology pair: a difficulty index and a benefit index. To this end, we need to make technical comparisons that require deep knowledge of the underlying tools. We argue that LLMs are well suited to this task: trained on a broad corpus of software usage documentation and code, they not only possess deep knowledge of this software but also enable us to apply a single, consistent rubric across all software stacks. Nonetheless, LLM-based measures must be treated with caution, as they can be sensitive to model choice, which can affect downstream empirical conclusions (Yin et al., 2026). To validate these measurements, we compare the results across model providers and also validate them against independent human survey ratings.[5]

For each legacy–modern pair, we ask the LLM (ChatGPT 5.4) to evaluate how difficult it would be for a worker with two years of experience in the legacy technology to independently complete routine productive tasks with the modern technology. The model rates the difficulty of this move based on the conceptual shift required and the new tooling or environment burden. We standardize this direct difficulty score across the 15 technology pairs. Higher values therefore indicate modern technologies that require a larger conceptual and tooling adjustment from the legacy starting point. Appendix Figure 11 shows the distributions of the standardized difficulty and adoption-benefit indices across the fifteen technology pairs.

To construct the benefit index, we ask the model to rate the practical advantage of the modern technology over the legacy technology along six dimensions: code simplicity, runtime efficiency, feasibility at production scale, maintainability, safety and reliability, and

[5]Our baseline ratings use an OpenAI model. As described below, we cross-check these against ratings from another model provider, as well as independent human survey ratings, to assess robustness.

flexibility.[6] Unlike the difficulty measure, benefits can be negative if the new technology is inferior to the legacy technology in certain aspects. The score thus ranges from -5 to 5. We average these six scores and standardize the resulting pair-level index across the fifteen technologies. Higher values thus indicate technologies for which the modern stack offers larger practical benefits relative to the legacy stack.

The mechanism specification is a triple-difference extension of the event study. The mechanism test uses the continuous standardized difficulty and adoption-benefit indices, denoted D_j and B_j. We estimate

$$
\begin{aligned}
Y_{cjt} = \alpha_{cj} + \gamma_{j\tau} + \lambda_{s(c)t} + \sum_{\tau \neq -1} \beta_\tau \left[\mathbf{1}\{\tau_{jt} = \tau\} \times \text{OldShare}_c^{\text{col}} \right] \\
+ \sum_{\tau \neq -1} \theta_\tau \left[\mathbf{1}\{\tau_{jt} = \tau\} \times \text{OldShare}_c^{\text{col}} \times D_j \right] + \sum_{\tau \neq -1} \lambda_\tau \left[\mathbf{1}\{\tau_{jt} = \tau\} \times \text{OldShare}_c^{\text{col}} \times B_j \right] \\
+ X'_{ct}\delta + \varepsilon_{cjt}. \qquad (3)
\end{aligned}
$$

The endogenous age-composition terms are instrumented using the aforementioned Bartik instrument. Specifically, each interaction involving the local old-worker share is instrumented with the corresponding interaction using the predicted old-worker share. The coefficients θ_τ and λ_τ measure the change in the old-workforce adoption penalty associated with one standard deviation increase in the difficulty and benefit indices, respectively. Figure 9 reports the implied old-workforce effect at -1 and $+1$ standard deviation of the difficulty index, holding the benefit index at its mean. Appendix Figure 12 reports the corresponding comparison for the benefit index. The outcome is modern-technology ads as a share of all job ads. The specification uses the same preferred controls as the main all-ad-share event study.

The triple-difference results align with the difficulty hypothesis. For modern ads as a share of all job ads, the implied aging penalty is more negative for technologies one standard deviation above the mean difficulty index than for technologies one standard deviation below it, especially in the later post-release years. Higher-benefit technologies also show a steeper old-workforce adoption penalty at year 4, but we interpret this pattern more cautiously because the benefit interaction is not robust to our permutation test robustness check.

[6] These dimensions were chosen based on focus-group interviews with technology professionals about the main practical margins along which new software tools improve on older ones.

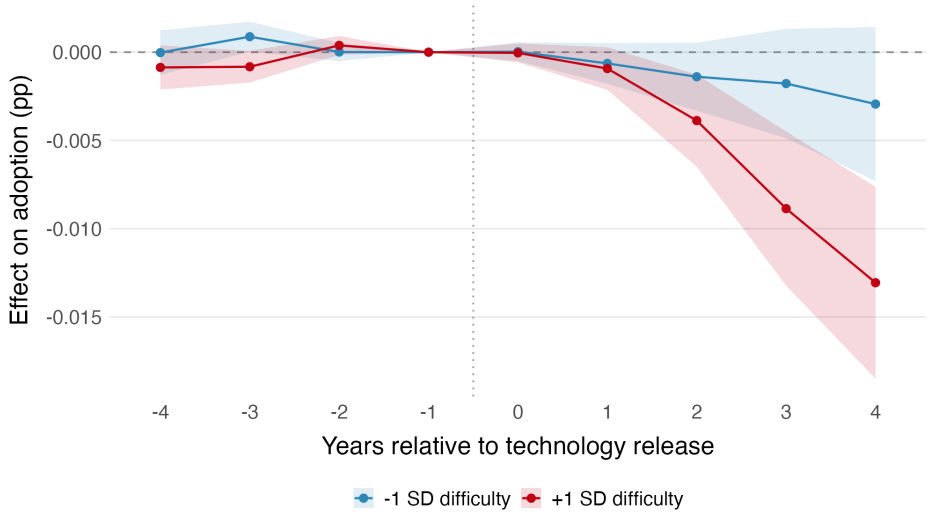


Figure 9: Implied old-workforce adoption penalties at −1 and +1 standard deviation of the difficulty index

Appendix Table 9 reports two robustness checks for the heterogeneity exercise. First, since our technology sample is quite small, we conduct a permutation test to check whether the observed interactions are unusually large relative to placebo assignments of the same technology-level scores. We randomly permute the paired difficulty and benefit labels across technologies 1,000 times and re-estimate the joint triple-difference specification for each permutation. Second, we want to ensure that the results are not driven by a single technology outlier by running a leave-one-technology-out test. The difficulty-index pattern passes both tests. In the permutation test, the year-4 difficulty interaction has a two-sided permutation p-value of 0.006, and the coefficient remains negative in all fifteen leave-one-technology-out estimates. The benefit interaction, on the other hand, is less robust: its permutation p-values are large, and its leave-one-out coefficients are also highly unstable.

An alternative interpretation of the adoption gap is that it reflects adoption frictions other than learning costs, such as information or risk. These stories, however, cannot explain all of our empirics. A pure information friction or pure risk story does not naturally predict that the age gradient should be steeper specifically for technologies that require larger conceptual and tooling transitions from the legacy stack, nor it is consistent with a progressively larger gap, given that information friction or uncertainty should diminish with time.

Lastly, it is possible that local age structure is highly correlated with the age structure of managers within firms, and that it is the age structure of managers that drives the results rather than that of the workforce. To address this alternative explanation, we run an additional regression controlling for the interaction between event time and the share of older managers in the CZ, while retaining the full set of preferred controls and state-by-year fixed

effects. The results, reported in Appendix Figure 21, show that the manager age-structure coefficients are close to zero and statistically insignificant, while the workforce coefficients remain robust.

6.1 Validation of LLM Measures

We check whether the LLM estimates are stable across model providers and whether they align with independent human judgments. ChatGPT and Claude (Sonnet 4.6) scores produce similar technology rankings: the difficulty measures have Pearson and Spearman correlations of 0.93 and 0.92, and the benefit measures have correlations of 0.86 and 0.86.

In addition to the cross-provider comparison, we also assess the indices using an ongoing survey of computer scientists, IT professionals, and students. We sent out an online Qualtrics survey to the Yale Computer Science community as well as IT communities on social media. We also exclude observations flagged for suspiciously short survey completion times, repetitive responses, or belonging to IP or browser address clusters that have a high incidence of observations with low CAPTCHA scores. More details about data screening criteria can be found in Appendix C.

After removing unusable observations, we are left with 710 completed surveys. Of those, 438 surveys have difficulty and benefit ratings for at least one legacy-modern pair. When rating a technology pair, survey respondents are first asked to rate the extent to which the modern technologies in a pair required substantially new thinking, tools and knowledge of new technical ecosystems compared to their legacy counterparts (i.e. , the difficulty rating). Respondents are presented with 5 ordinal choices with text descriptions evenly spaced and coded between 0 to 5. They are then asked to rate the modern technologies on the six characteristics rated by the LLM relative to their legacy counterparts (i.e. , the benefit ratings), choosing from 5 ordinal responses between much less beneficial to much more beneficial. We average respondents' available ratings of each index for each technology pair, center these indices for each respondent by their mean index rating over all technology pairs that they rated and then average centered observations for each technology pair.

The LLM-constructed difficulty index agrees with respondent-centered human ratings. Among the twelve pairs with at least 20 human ratings, the Pearson and Spearman correlations are 0.75 and 0.75, respectively. Agreement is weaker for the benefit index, with corresponding correlations of 0.21 and 0.28. The disagreement, however, does not alter our main triple-difference conclusion for the difficulty index, as shown in Appendix D. We therefore view the survey as providing stronger support for the difficulty index and interpret the benefit measure more cautiously.

7 Model

Our empirical results are consistent with three mechanisms: a shorter remaining work horizon, greater loss of accumulated legacy experience upon switching, and a higher learning burden due to reduced learning efficiency.

To quantify these three mechanisms, we construct a parsimonious model that links workers' retraining decisions with firms' adoption decisions.

7.1 Environment

This paper considers a single technology cycle: an incumbent technology already in use, and a frontier technology that arrives to replace it.

Workers enter the labor market at age zero and retire at age T. Thus, a worker of age a has $T-a$ working years remaining. Workers initially use the incumbent technology, denoted I, and accumulate experience with it.

At time $t = 0$, a superior frontier technology, denoted F, with difficulty ϕ, arrives. Workers choose whether to learn the new technology or stay with the incumbent. The decision is irreversible. New workers face the same decision at entry.

Frontier and incumbent technologies are perfect substitutes. The firms' production function takes the form

$$Y_t = A\left(L_{I,t} + L_{F,t}\right)^{\theta}, \tag{4}$$

where $L_{I,t}$ and $L_{F,t}$ denote effective labor employed with the incumbent and frontier technologies at calendar time t, respectively, $A > 0$ is aggregate productivity, and $\theta > 0$ is the output elasticity of effective labor. Thus, under competition, wages per unit of effective labor are

$$w_{I,t} = w_{F,t} = w_t = A\theta\left(L_{I,t} + L_{F,t}\right)^{\theta-1}. \tag{5}$$

7.1.1 Workers' Problem

Let $a \in [0, T)$ denote a worker's age when choosing between the two technologies, and let $s \in [0, T-a]$ denote time since that choice. Each worker has a learning efficiency $\mu(a)$, which measures experience gained per year and declines with age:

$$\mu(0) = 1, \qquad \mu(a) > 0, \qquad \mu'(a) \leq 0.$$

If the worker stays with the incumbent technology, accumulated experience is:

$$X_I(a,s) = \int_0^{a+s} \mu(v)\, dv. \tag{6}$$

If the worker switches to the frontier technology, accumulated experience is lost because experience is technology-specific:

$$X_F(a,s;\phi) = \frac{1}{1+\phi} \int_a^{a+s} \mu(v)\, dv, \qquad \phi \geq 0. \tag{7}$$

The more difficult the new technology, the slower experience accumulates and the later the worker receives its benefits. The learning rate is

$$\frac{\partial X_F(a,s;\phi)}{\partial s} = \frac{\mu(a+s)}{1+\phi}.$$

Productivity depends on experience through $K(X)$, with $K(0) = 1$, $K'(X) > 0$, and $K''(X) \leq 0$. Each worker has an aptitude $z_i > 0$ for the frontier technology that makes the worker more or less productive with it. The frontier technology provides a common productivity gain $\alpha > 1$. Workers have perfect foresight about the calendar-time wage path $\{w_t\}_{t\geq 0}$ and take it as given when choosing their technology. For a worker choosing at $t = 0$, earnings s years after the choice are

$$y_{I,i}(a,s) = w_s K\Big(X_I(a,s)\Big), \qquad y_{F,i}(a,s;\phi) = w_s \alpha z_i K\Big(X_F(a,s;\phi)\Big).$$

Workers also face a one-time nonpecuniary learning burden $c(a,\phi)$ that increases with age and the difficulty of the technology. Measured in units of discounted earnings, this cost satisfies

$$c(a,\phi) \geq 0, \qquad \frac{\partial c(a,\phi)}{\partial a} \geq 0, \qquad \frac{\partial c(a,\phi)}{\partial \phi} \geq 0. \tag{8}$$

The worker compares discounted earnings from staying with the incumbent and switching to the frontier, net of the learning cost. With discount rate $r > 0$, the value of staying is

$$V_I(a) = \int_0^{T-a} e^{-rs} w_s K\Big(X_I(a,s)\Big)\, ds, \tag{9}$$

and the value of switching is

$$V_F(a,z_i;\phi) = \alpha z_i \int_0^{T-a} e^{-rs} w_s K\Big(X_F(a,s;\phi)\Big)\, ds - c(a,\phi). \tag{10}$$

Let $B_I(a) \equiv V_I(a)$ and

$$B_F(a;\phi) = \int_0^{T-a} e^{-rs} w_s K\Big(X_F(a,s;\phi)\Big)\, ds.$$

Since $B_F(a;\phi) > 0$ for $a < T$, this comparison gives an aptitude threshold above which the worker adopts:

$$V_F(a, z_i;\phi) \geq V_I(a) \quad \Longleftrightarrow \quad z_i \geq z^*(a;\phi) \equiv \frac{B_I(a) + c(a,\phi)}{\alpha B_F(a;\phi)}. \tag{11}$$

The worker adopts when discounted frontier earnings are sufficient to cover the earnings given up under the incumbent technology and the learning cost. New entrants face the same decision at $a = 0$, with no accumulated incumbent experience.

There are three forces that make adoption less attractive for older workers. First, the shorter horizon means that they will have less time to recoup the investment cost, as in Ben-Porath (1967). Second, the learning burden increases with age due to reduced learning efficiency. Third, they have to give up more incumbent experience to adopt the frontier technology.

On the other hand, greater technology difficulty may reduce the value of adoption more for younger workers because they remain employed longer. As illustrated in Figure 10, frontier earnings may initially exceed incumbent earnings, but slow accumulation of frontier experience may subsequently reverse this relationship. Older workers may retire before incumbent earnings overtake frontier earnings, whereas younger workers bear the resulting earnings losses. If these later losses are sufficiently large, even after discounting, the value of adoption may be lower for younger workers. This scenario requires both a sufficiently high aptitude z_i or frontier advantage α to generate an immediate productivity gain from switching, and extremely high wages in the additional working years for the subsequent losses to outweigh that gain. Under single crossing assumption where frontier earnings path can only cross incumbent earnings once from below, we can show that younger workers are more likely to adopt than older workers.

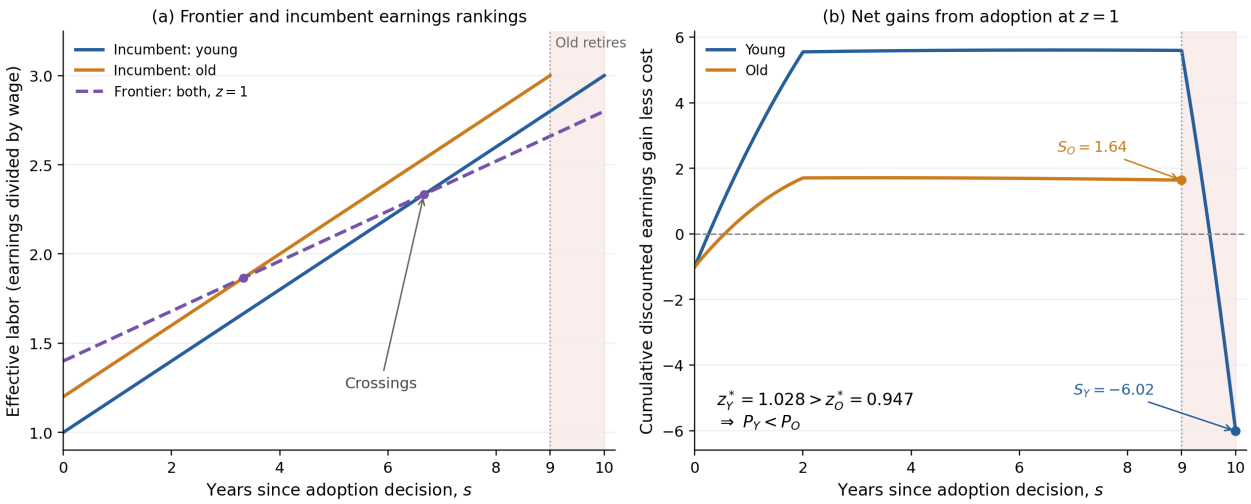


Figure 10: An illustrative reversal of the age ranking in adoption

Notes: Panel (a) shows earnings profiles for a young and old worker using I and F. Panel (b) shows cumulative discounted earnings gains net of the learning cost. Net gains are positive for the older worker and negative for the younger worker.

7.1.2 Aggregation and Equilibrium

Effective labor under each technology is the sum of the productivity supplied by workers choosing that technology. Incumbent workers supply $K(X_I)$, while frontier workers supply $\alpha z K(X_F)$. Aggregation includes initial workers who have not retired and subsequent entrants who are still working, taking account of their age and aptitude. Future entrants choose between the two technologies at entry, on the basis of earnings over their remaining working lives.

Given the initial age distribution and the sequence of entering cohorts, a perfect-foresight equilibrium consists of a wage path, workers' technology choices, and effective labor under each technology. Workers maximize discounted lifetime earnings given the anticipated wage path; effective labor supplied by their choices equals firms' demand under each technology; and firms pay the marginal product of effective labor in equation (5) at every date. The wage path workers anticipate is the one generated by these choices. We solve for equilibrium separately in each commuting zone. Appendix K gives the cohort aggregation formulas and numerical solution procedure.

7.2 Calibration

The calibration treats each commuting zone as a separate labor market. Initial age distributions are taken from the pooled 2010–2011 ACS sample of college-educated workers ages 20–69, with each initial workforce normalized to one. Workers enter at age 20 and retire at

age 70, so $T = 50$. We set the annual discount rate to 4% and the annual growth rate of entering cohorts to 0.7%, matching observed US population growth over the sample period.

The functions governing learning efficiency, the return to experience, and the learning burden are respectively:

$$\begin{aligned} \mu(a) &= \frac{1}{1+\lambda_\chi a}, \\ K(X) &= \exp\left[K_{\max}\left(1-e^{-\rho_K X}\right)\right], \\ c(a,\phi) &= \frac{\phi}{\mu(a)}. \end{aligned} \tag{12}$$

With this specification, declining learning efficiency increases the learning burden as well as the time required to accumulate experience. We assume that frontier aptitude is lognormally distributed, with mean one:

$$z_i = \exp\left(\varepsilon_i - \frac{\sigma_\varepsilon^2}{2}\right), \qquad \varepsilon_i \sim N(0,\sigma_\varepsilon^2). \tag{13}$$

We set $K_{\max} = 0.883$ and $\rho_K = 0.131$, using the fit to the 2011 ACS hourly-wage profile of college-educated computer workers. To obtain the age profile of learning efficiency, we use problem-solving scores from the Programme for International Assessment of Adult Competencies (PIAAC), a survey of adult problem-solving skills in a technology-rich environment. [7] We regress standardized problem-solving scores on age, controlling for sex, education, and survey wave. Using the estimated age gradient, we set $\lambda_\chi = 0.0171$ in $\mu(a) = 1/[1+\lambda_\chi a]$, normalizing learning efficiency at age 20 to one. This specification implies that a worker aged 65 accumulates about 56% as much effective experience per year as a worker aged 20. Appendix J reports the PIAAC profile and the ACS experience–earnings comparison.

The remaining parameters, α, σ_ε, and ϕ, are chosen to match three moments measured at empirical event year 4, which we treat as the date of the model's initial adoption decision. The first is the frontier posting share, 24.94%, measured as the share of postings mentioning the frontier technology among postings mentioning either the frontier or its matched legacy technology. This moment helps pin down the frontier advantage and the difficulty as it restricts the overall return to adoption relative to its cost. The second is the event-year-4 age coefficient, -0.230, from the weighted matched-pair IV event study reported in Figure 7. We calculate its model counterpart by regressing simulated adoption on the share of workers aged 50 and above across commuting zones. This moment governs how strongly

[7] PIAAC assesses adults' problem-solving skills with digital tools. Tasks involve simulated computer applications, such as navigating websites, managing emails, which often required respondents to coordinate several steps across applications. The assessment aims to capture the joint use of reasoning and computer skills.

adoption costs differ with age. The third moment is the frontier log-earnings premium, 0.0779, estimated from Lightcast posted wages with technology-pair-by-commuting-zone-by-occupation-by-seniority fixed effects. This moment helps identify both the productivity advantage of the frontier and the distribution of workers' aptitude. To validate the model, we check its fit for an untargeted moment – the standard deviation of frontier residual log earnings.

For each candidate set of parameters, we calculate age-specific adoption cutoffs and aggregate adoption and earnings using the observed local age distributions. We choose the parameters that minimize the sum of squared differences between model and data moments. We set $\theta = 2/3$, so production exhibits diminishing returns. Worker choices and the competitive wage path are determined jointly. We normalize A so that the wage per unit of effective labor equals one when effective labor is at its national reference level with all workers using the incumbent technology.

Table 3 reports the estimates and the corresponding model moments. The model matches the three targets and also approximates the untargeted moment reasonably well: it produces a standard deviation of 0.251, compared with 0.232 in the data.

Table 3: Calibration parameters and fit

	Data	Model
Estimated parameters		
Frontier productivity advantage α	—	1.517
Aptitude dispersion σ_ε	—	0.436
Technology difficulty ϕ	—	1.579
Targeted moments		
Frontier posting share (%)	24.94	24.94
Cross-market age coefficient	−0.230	−0.230
Frontier log-earnings coefficient	0.0779	0.0779
Untargeted moment		
Frontier residual log-earnings SD	0.2323	0.2514

Alternative calibration without experience loss

One may be curious as to whether a model with only cognitive decline and shorter horizons can explain the data, or question the mapping from PIAAC test scores to learning efficiency. We thus allow the learning efficiency gradient to vary and re-estimate the remaining parameters to match the targeted moments. The resulting cognitive decline parameter λ_χ implies that learning efficiency at age 65 is only 15% of its age-20 level, compared with 56% under our current specification. Additionally, even with such a steep cognitive-decline gradient, the

untargeted earnings dispersion is still not well-matched. Appendix L reports the exercise.

7.3 Channels Decomposition

We next use the calibrated model to examine how the age difference in adoption changes when experience loss, reduced learning efficiency, or differences in remaining work horizons are removed separately. First, we shut down the experience loss channel by allowing workers to retain their accumulated incumbent experience when switching. Frontier experience then starts at $X_I(a, 0)$, while difficulty continues to slow subsequent learning. Second, we shut down cognitive decline by equalizing learning efficiency across ages by setting $\mu(a) = 1$ at every age. Experience then accumulates at the same rate across ages, and the learning burden is ϕ for all workers.Third, we give every worker a remaining work horizon of $H = 30$ years, replacing $T-a$ with H in both earnings integrals. while leaving the other two channels intact. [8]

Table 4 shows that the shutdown of the experience loss channel substantially reduces the age difference in adoption. The difference falls from 32.3 to 3.9 percentage points, a reduction of about 88 percent. Equalizing remaining horizons at 30 years reduces the gap to 18.6 percentage points, a reduction of 42.5 percent, indicating moderate importance of the channel. Giving all workers 20 or 40 remaining years reduces the gap by 41.3 or 43.5 percent, respectively. Meanwhile, shutting down cognitive decline barely has any effect on the gap, suggesting that it plays a minor role in explaining adoption gap. Among these separate comparisons, experience loss is the most important mechanism driving adoption gap.

Table 4: Adoption with each channel removed separately

	All workers (%)	Age 20–34 (%)	Age 50–69 (%)	Old–young gap (pp)
Full model	24.9	41.8	9.5	−32.3
No experience loss	68.3	64.9	61.1	−3.9
Equal learning efficiency	30.8	46.0	14.5	−31.4
Equal remaining horizons (30 years)	26.9	38.5	19.9	−18.6

Notes: Aggregate adoption uses observed age shares and posting weights across commuting zones. The gap is older minus younger adoption. Channels are removed one at a time.

7.4 Counterfactuals

Workforce aging can result from lower fertility, which slows the inflow of young workers, or from later retirement. Both increase the share of older workers, but their implications for

[8] We also try setting $H = 40$ and $H = 20$, and the results remain largely unchanged.

adoption differ. Lower fertility increases the share of older workers but only affects worker choices through the wage feedback. On the other hand, later retirement not only changes the age distribution but also changes workers' behavior directly by increasing the return to skill investment. It is thus interesting to understand the net effect of each source of aging.

We hold the calibrated parameters fixed and compare economies at the same date of frontier arrival. Let g denote the annual growth rate of successive entering cohorts and R the retirement age in calendar years. Initial age shares are given by

$$h_{g,R}(a) = \frac{(1+g)^{-a}}{\sum_{j=0}^{R-21}(1+g)^{-j}}, \qquad a = 0, \ldots, R-21, \tag{14}$$

where a is years since entry at age 20. Our benchmark economy has zero population growth rate, i.e g = 0 and retirement age set to 70.

For easier interpretation, we can translate the population growth into fertility rate since the total fertility rate is the policy-relevant. If we assume zero net migration, a fixed low-mortality schedule, and a generation length of 30 years, then in the long run, a cohort growth rate g then corresponds to a total fertility rate of approximately TFR $\approx 2.1(1+g)^{30}$. An economy with g = 0 corresponds to TFR at replacement rate (2.1), while annual contraction and growth of 1% correspond to total fertility rates of roughly 1.6 and 2.8.

In the reference economy, fertility is at replacement level and workers retire at age 70. Initial frontier adoption is 23.39%.

Table 5: Demographic counterfactuals relative to replacement-level fertility and retirement at age 70

	Change in adoption
TFR = 1.6	−7.5%
TFR = 2.8	+7.4%
Retirement at 75, TFR = 2.1	+0.45%

Notes: Fertility rates are the illustrative long-run equivalents of annual cohort growth rates $g = -1\%, 0, +1\%$ under zero net migration, a fixed low-mortality schedule, and a 30-year generation length, using TFR $\approx 2.1(1+g)^{30}$.

Table 5 shows that population growth matters substantially for adoption. Relative to an economy at replacement-level fertility, a total fertility rate of 1.6 lowers aggregate adoption by 7.5%, while a rate of 2.8 raises it by 7.4%. By contrast, raising the retirement age by five years has little effect on aggregate adoption because longer remaining work horizons encourage adoption, while the higher share of older workers lowers it (Table 6).

Table 6: Decomposition of the effect of raising the retirement age from 70 to 75

Component	Change in adoption (%)
Longer work horizon, reference wages and age distribution	+9.62
Wage response, reference age distribution	−0.04
Change in age distribution	−9.13
Total effect	+0.45

Notes: Entries are changes expressed as percentages of reference adoption. Fertility is fixed at replacement rate (g = 0). The decomposition first changes the work horizon while keeping wages and age weights fixed, then substitutes the counterfactual wage path, and finally substitutes the counterfactual age distribution.

8 Conclusion

Workforce aging hampers technology adoption. Within four years of a new technology's release, there emerges a divergence between young and old commuting zones that widens over time. This pattern is robust across multiple proxies of adoption, and is driven by a differential adoption cost between young and old commuting zones.

Building on the empirical evidence, we develop a parsimonious model linking worker behavior to firms' technology adoption decisions, which separates three distinct mechanisms driving the difference between old and young workers: shorter remaining work horizons, experience loss, and reduced learning efficiency. The calibrated model shows that demographic policies can affect technology adoption in different—even opposing—ways. Falling fertility, which slows the inflow of younger workers, lowers adoption by shifting the workforce age distribution older: holding net migration at zero, moving from replacement-level fertility to a total fertility rate of 1.6 reduces aggregate adoption by 7.5 percent. Raising the retirement age has a more subtle effect: it makes the measured workforce older on average, but also lengthens workers' remaining horizon, which can increase adoption. Given how common aging-related policies are, and how important technology is as a driver of growth, policymakers need to think carefully before implementing policies that shift a region's age composition, since such policies do not all affect technology adoption in the same way.

Several limitations remain. We use job postings requiring software skills as a proxy for direct technology use inside firms, and our analysis focuses on major IT software technologies as canonical examples of skill-complementary technologies. While we believe the analysis extends naturally to all skill-complementary technologies, this remains to be verified. The model also abstracts from general equilibrium responses and focuses on retraining and adoption following the arrival of one new technology. Future work should extend the framework to

successive arrivals. Another promising direction is studying how generative AI by changing the nature of learning costs themselves may reshape these dynamics. More broadly, the open question is whether aging economies simply need different policy levers to sustain adoption, or whether adoption capacity, once lost to an aging workforce, is difficult to recover.

Appendix

A Technology Sample

Table 7: Modern technologies in the main sample and their legacy comparisons

Modern technology	**Release**	**Category**	**Legacy comparison**
Apache Spark	2014	Data processing	Hadoop MapReduce
React	2013	Front-end framework	JavaScript / jQuery
Docker	2013	DevOps / deployment	Virtual machines / on-prem deployment
Rust	2015	Programming language	C/C++
Keras	2015	Machine learning	Theano
Apache Hive	2012	Data query engine	MapReduce / SQL workflows
Go	2012	Programming language	C/C++
TypeScript	2012	Programming language	JavaScript
Angular	2014	Front-end framework	JavaScript / jQuery
Vue.js	2014	Front-end framework	JavaScript / jQuery
Swift	2014	Programming language	Objective-C
Kubernetes	2015	Container orchestration	Virtual machines / manual deployment
Kotlin	2016	Programming language	Java
Elasticsearch	2014	Search / indexing	Apache Solr / Lucene
Ansible	2013	Infrastructure automation	Puppet

B Index Distributions

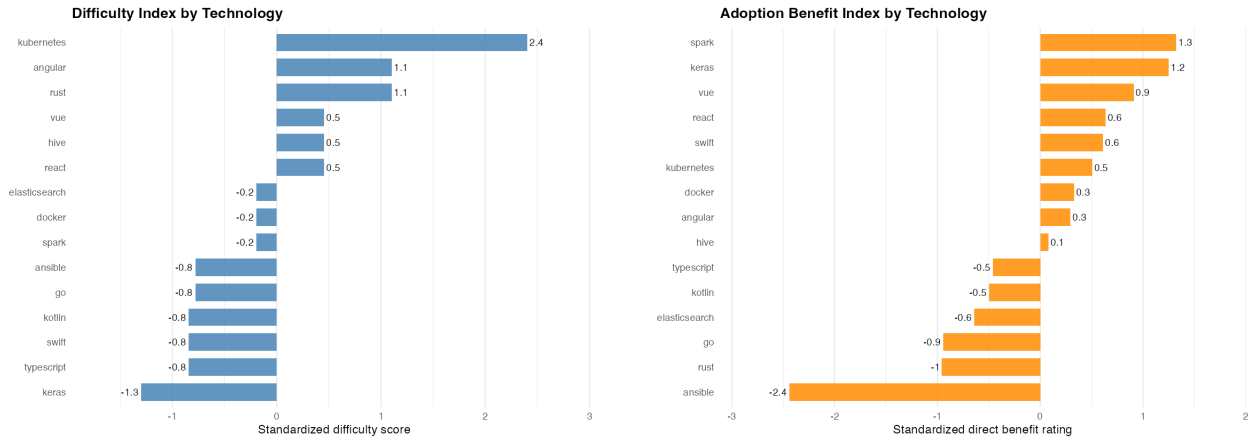


Figure 11: Distributions of the difficulty and adoption-benefit indices

C Stability and Validity of LLM Measures

One concern is that LLM-based measures are sensitive to model choice or model vintage and therefore not reliable enough on their own. Yin et al. (2026) show that LLM-based occupational AI-exposure scores vary substantially across frontier models, and that these differences can change downstream empirical conclusions. To address this concern, we check the robustness of our LLM-based technology rankings across frontier model providers, OpenAI and Anthropic, and validate the resulting indices against independent human survey ratings.

The provider comparison is reassuring. For the difficulty index, OpenAI and Claude are strongly aligned across the fifteen technology pairs, with a Pearson correlation of 0.93 and a Spearman rank correlation of 0.92. The benefit indices are also strongly correlated, with a Pearson correlation of 0.86 and a Spearman rank correlation of 0.86. Thus, while provider choice changes some pair rankings, the overall ordering of technologies is stable enough that the heterogeneity results are unlikely to be driven by arbitrary LLM noise.

Nonetheless, we might still be worried that the training materials or algorithms contain shared biases. We therefore assess the LLM-based indices using an online survey of IT professionals and students. The survey was distributed to the Yale Computer Science community and IT communities on Reddit, Discord, and LinkedIn. Data collection is ongoing, and respondents include students, industry professionals, and professors.

To ensure response quality, we exclude incomplete responses as well as responses suspected to be automated. We first remove any response with a CAPTCHA score, as de-

termined by Qualtrics, of 0.6 or below out of 1.[9] We next exclude responses completed suspiciously quickly, defined as those with at least 16 answered rating items and a total survey duration averaging under two seconds per answered rating item. We then drop exact duplicates, along with near-identical responses completed within one hour of one another. Two responses are considered near-identical if they share at least 24 of the 127 possible non-missing rating items, covering at least 80% of each respondent's answered items, and agree on at least 95% of those shared ratings. Finally, we remove respondents with repeated answer patterns—those who give the same answers across different technology pairs—as these may indicate inattentive responding.

We additionally remove bursts of low-quality responses on suspicion of automation. We flag all 60-second intervals containing at least five survey starts, at least half of which have observed CAPTCHA scores below 0.5, and drop every response within such an interval. After these exclusions, our sample contains 710 completed responses, of which 438 provide usable difficulty and benefit ratings for at least one legacy–modern technology pair.

For each pair that a respondent can assess, we ask how much the modern technology requires them to think in a fundamentally new way relative to the legacy technology and how much new tooling and ecosystem knowledge must be learned. These two questions are intended to parallel the LLM difficulty prompt. Each component is scored on a scale of 0 to 5. For benefits, respondents also evaluate the six components assessed by the LLM, from "Much less beneficial" to "Much more beneficial." We then take the average of the available items within each index and respondent–technology pair. Because respondents have their own scale of evaluation given their knowledge of different technology pairs, we subtract each respondent's mean across the pairs they rated and then average these centered scores within each technology pair.

The difficulty index is positively related to human judgments. Among the twelve pairs with at least 20 human ratings, the respondent-centered difficulty index has a Pearson correlation of 0.75 and a Spearman rank correlation of 0.75 with the LLM difficulty index. Both survey components show positive agreement with the overall LLM difficulty index: the corresponding correlations are 0.62 and 0.64 for new ways of thinking, and 0.79 and 0.73 for new tooling and ecosystem knowledge.

Agreement is weaker for benefits, which is unsurprising given that respondents may have different usage cases. Among the same twelve pairs, the respondent-centered benefit index has a Pearson correlation of 0.21 and a Spearman rank correlation of 0.28 with the LLM benefit index. When we examine the correlations of each component of the benefit index, we find variability. Human ratings of code efficiency (number of lines needed to complete

[9]Responses with missing CAPTCHA scores are retained.

the same task) are strongly correlated with the LLM's code-simplicity scores, with Pearson and Spearman correlations of 0.78 and 0.71, respectively, whereas the correlations for other components are much smaller. This may be because code simplicity is easier to assess consistently, whereas other components depend heavily on usage case and our sample size is too small to capture all usage cases. Since we do not know whether the LLM benefit index or the human-evaluated benefit index is more accurate, we re-ran the triple-difference regressions with the latter as a check, as presented in Appendix D. The main conclusion for the difficulty index remains robust. Overall, the survey provides stronger support for the difficulty index than for the benefit index. The Table 8 summarizes the correlations for pairs with at least 20 human ratings.

Table 8: Survey validation of LLM indices

Index	Minimum ratings	Pairs	Pearson	Spearman
Difficulty	20	12	0.75	0.75
Benefit	20	12	0.21	0.28

Notes: The sample contains 710 completed responses after the quality exclusions described above, with 438 respondents contributing to each index. We average available items within each respondent–technology pair, subtract each respondent's mean across their rated pairs, and average centered scores within pairs with equal respondent weights. Single-pair respondents contribute zero after centering. Correlations are calculated across the twelve pairs with at least 20 distinct respondents, giving each pair equal weight.

D Additional Triple-Difference Results

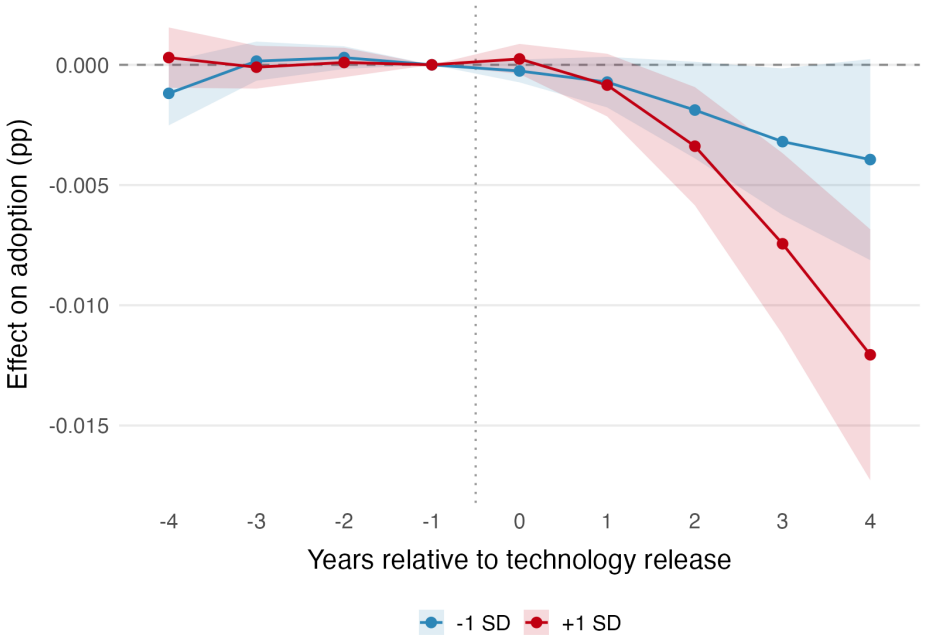
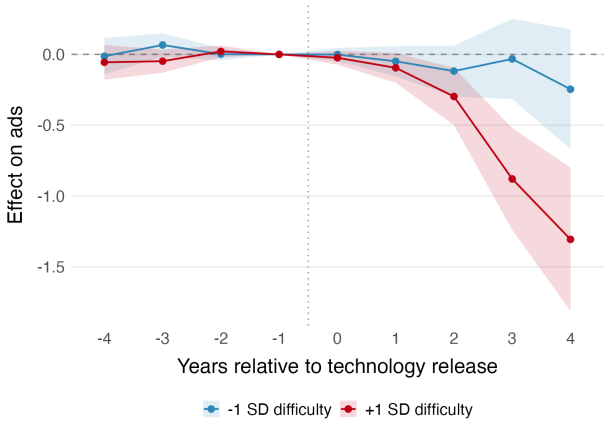
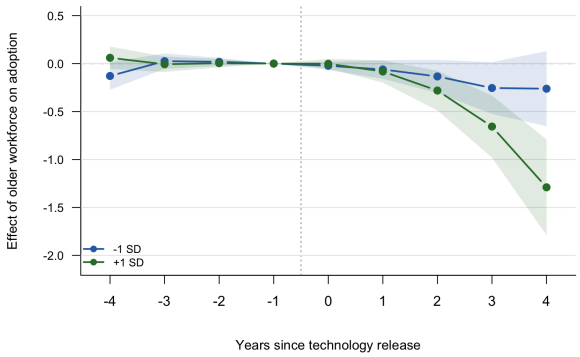


Figure 12: Implied old-workforce adoption penalties at −1 and +1 standard deviation of the benefit index, holding difficulty at its mean

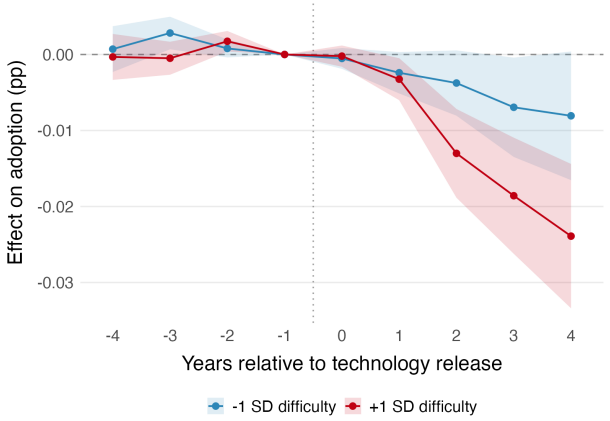
The main text reports continuous-index triple-difference estimates for the preferred outcome: new-technology ads as a share of all job ads. This appendix presents the corresponding estimates for two secondary outcomes: new-technology ads per 100,000 workers and the share of posting firms advertising for a new technology.



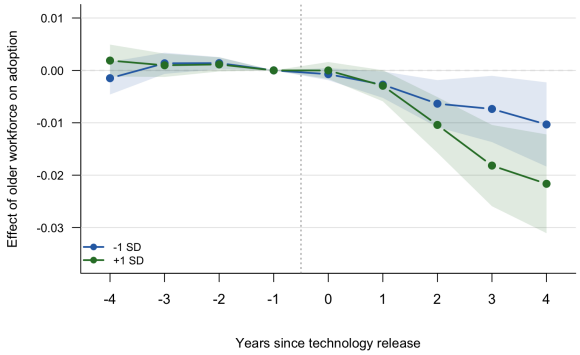
(a) Difficulty index, postings per worker



(b) Benefit, postings per worker



(c) Difficulty index, firm margin



(d) Benefit, firm margin

Figure 13: Implied old-workforce adoption penalties at -1 and $+1$ standard deviation of the continuous indices for secondary outcomes

Table 9 assesses sensitivity of the triple-difference results to permutation inference and the exclusion of individual technologies. For both posting-based outcomes, the year-4 difficulty interaction remains negative in every leave-one-technology-out specification, with permutation p-values of 0.006 and 0.014. The year-4 benefit interactions, by contrast, are not statistically distinguishable from zero under permutation inference. We therefore interpret the benefit results cautiously.

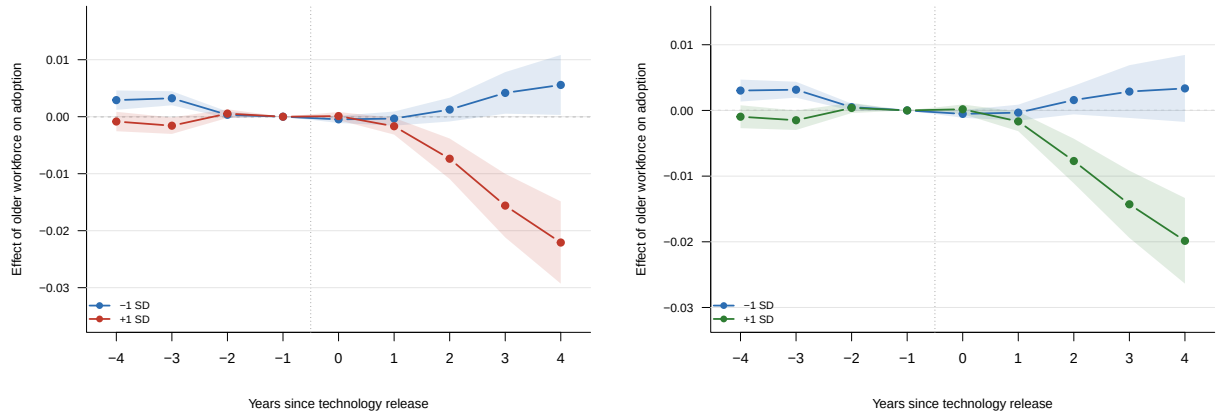
Table 9: Permutation and leave-one-out diagnostics for mechanism estimates

Outcome	Index	Year-4 coeff.	Perm. p	LOO range	Same sign
Modern share of all ads	Difficulty	-0.0228	0.006	[-0.0245, -0.0189]	15/15
Modern share of all ads	Benefit	-0.0010	0.916	[-0.0056, 0.0007]	13/15
Modern ads per 100,000 workers	Difficulty	-1.888	0.014	[-2.442, -1.511]	15/15
Modern ads per 100,000 workers	Benefit	-0.250	0.776	[-0.955, -0.088]	15/15

Notes: The table reports year-4 coefficients from the continuous-index IV triple-difference specification for the two modern-technology outcomes. "Perm. p" is the two-sided randomization p-value from 1,000 joint permutations of the technology-level difficulty and benefit scores across technologies. "LOO range" is the minimum and maximum coefficient from re-estimating the specification after omitting each technology in turn. "Same sign" reports the number of leave-one-out estimates with the same sign as the baseline estimate. Both outcomes are unweighted.

D.1 Human-Rated Benefits

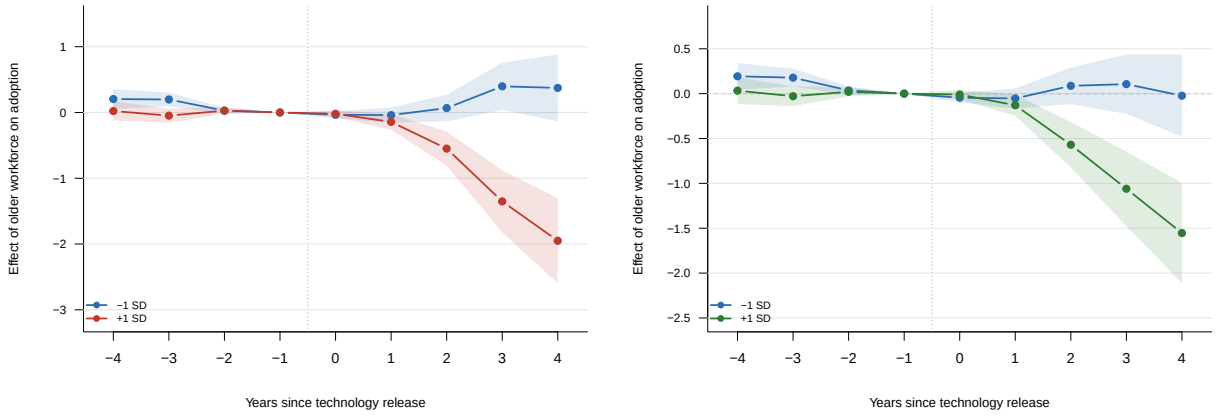
As a further robustness check, we replace the LLM-based benefit index with the respondent-centered human benefit index described in Appendix C, while retaining the LLM-based difficulty index. We restrict the analysis to the twelve technology pairs with at least 20 usable human benefit ratings. Figure 14 reports the human-benefit specification for new-technology ads as a share of all job ads. The difficulty gradient remains increasingly negative after release when we control for human-rated benefits.



(a) Difficulty index

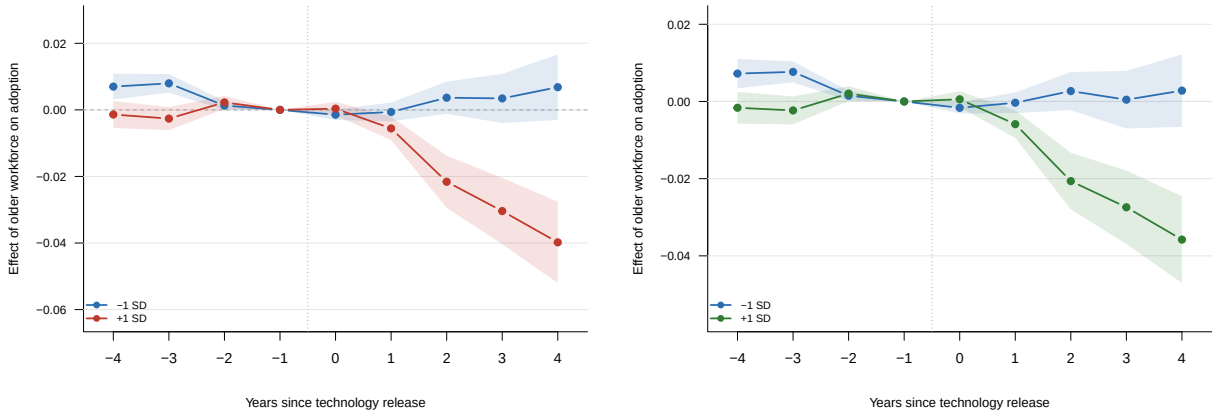
(b) Benefit index

Figure 14: Triple-difference estimates using human-rated benefits: new-technology share of all ads



(a) Difficulty index (b) Benefit index

Figure 15: Triple-difference estimates using human-rated benefits: new-technology ads per 100,000 workers



(a) Difficulty index (b) Benefit index

Figure 16: Triple-difference estimates using human-rated benefits: share of firms posting new technology

E Weighted Results

Figure 17 reports weighted versions of the all-ad-share, postings-per-worker, and firm-share outcomes. The weights vary by outcome: total job postings for all-ad shares, workforce size for postings per worker, and the number of posting firms for firm shares. Each weight is summed over the available pre-release years in the event window and held fixed over time, avoiding weights that depend on post-release changes in these denominators.

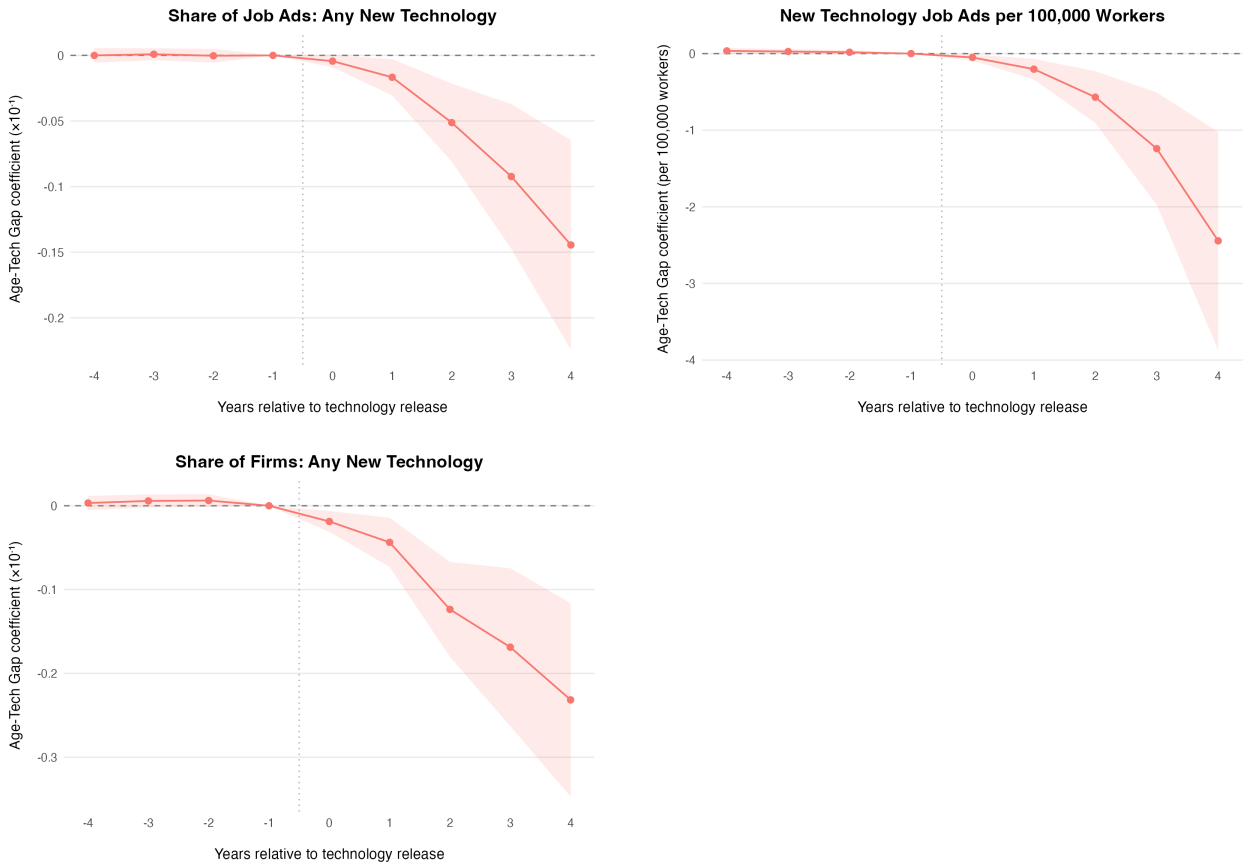
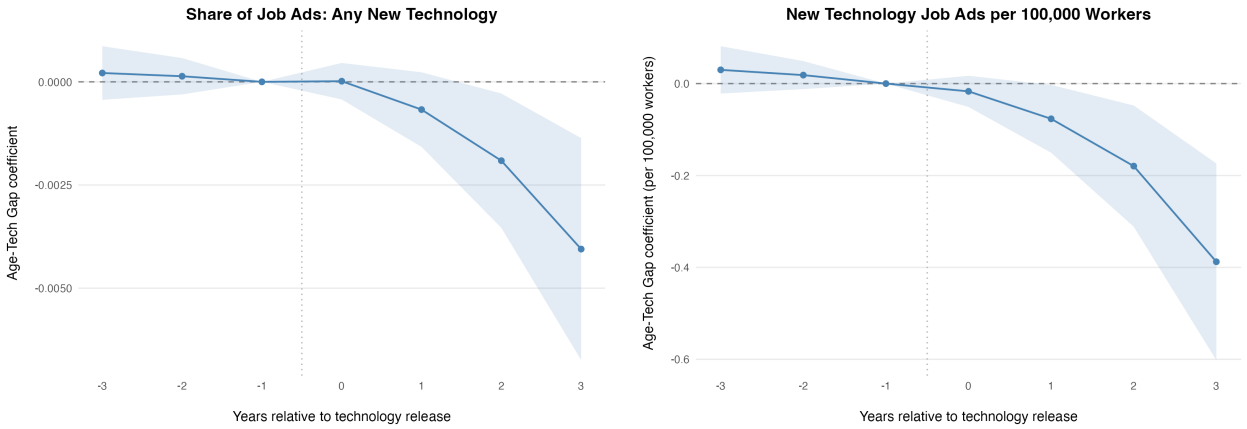


Figure 17: Weighted IV event-time estimates

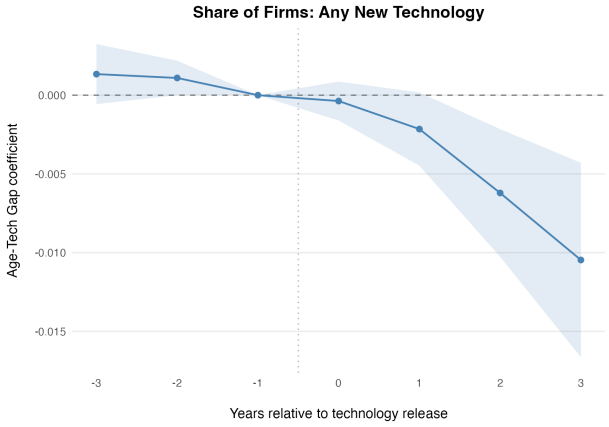
F Fixed Technology Composition

Because release years differ across technologies, longer event-time horizons are supported by different sets of technologies. Our estimates may therefore mix the actual causal effect with technology composition effect. As a common-support check, Figure 18 restricts the sample to the thirteen technologies observed in every event year from -3 through 3: Ansible, Docker, Elasticsearch, Hive, Keras, Kotlin, Kubernetes, React, Rust, Spark, Swift, TypeScript, and Vue. The qualitative results remain the same.

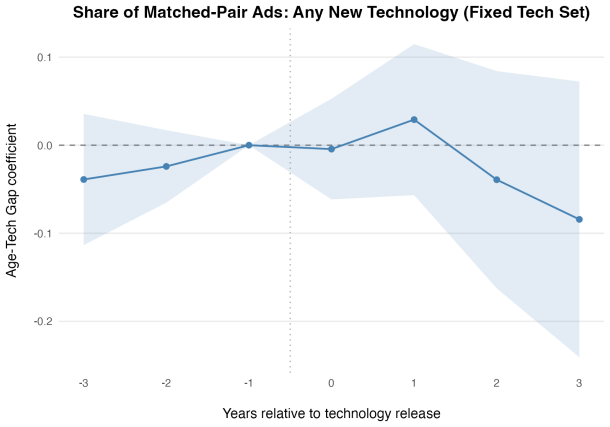


(a) Modern share of all ads

(b) Modern ads per 100,000 workers



(c) Firm margin



(d) Matched-pair modern share

Figure 18: Common-support IV event-study estimates using a fixed set of technologies from event time −3 through 3

G Other Robustness Checks

Figure 19 reports a robustness check that adds the number of local colleges and universities in 2000, interacted with event time, to the preferred IV specification. This addresses one of the main remaining balance concerns: older predicted-age CZs show some remaining difference in baseline university capacity and research infrastructure.

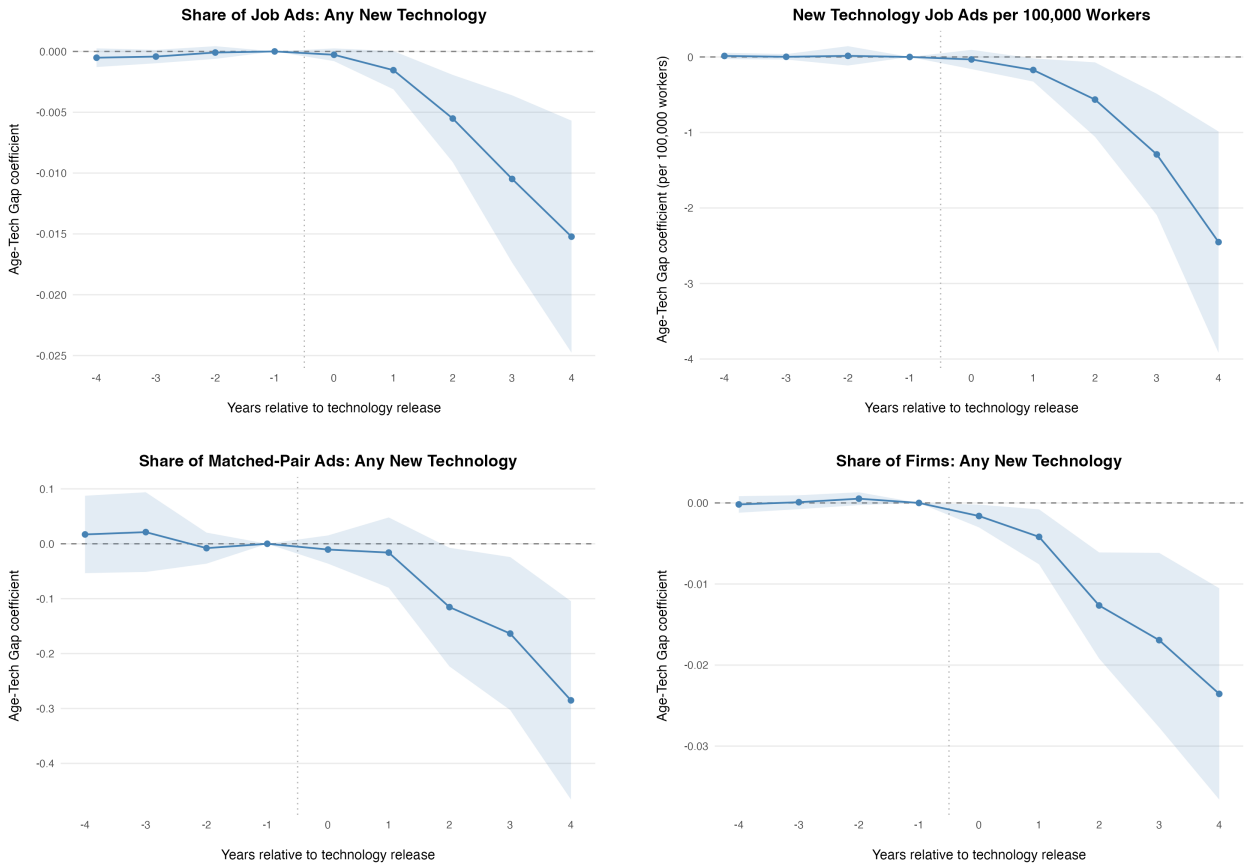


Figure 19: IV event-time estimates additionally controlling for university capacity

Figure 20 compares the estimated age gradient for the college-educated workforce with the corresponding gradient for workers with less than a high school degree across the all-ad-share, posting-count, and firm-level outcomes. The negative age gradient is concentrated in the college-educated margin, while the less-than-high-school coefficients are small and statistically insignificant. This pattern is consistent with the high-skill nature of the frontier software technologies studied here.

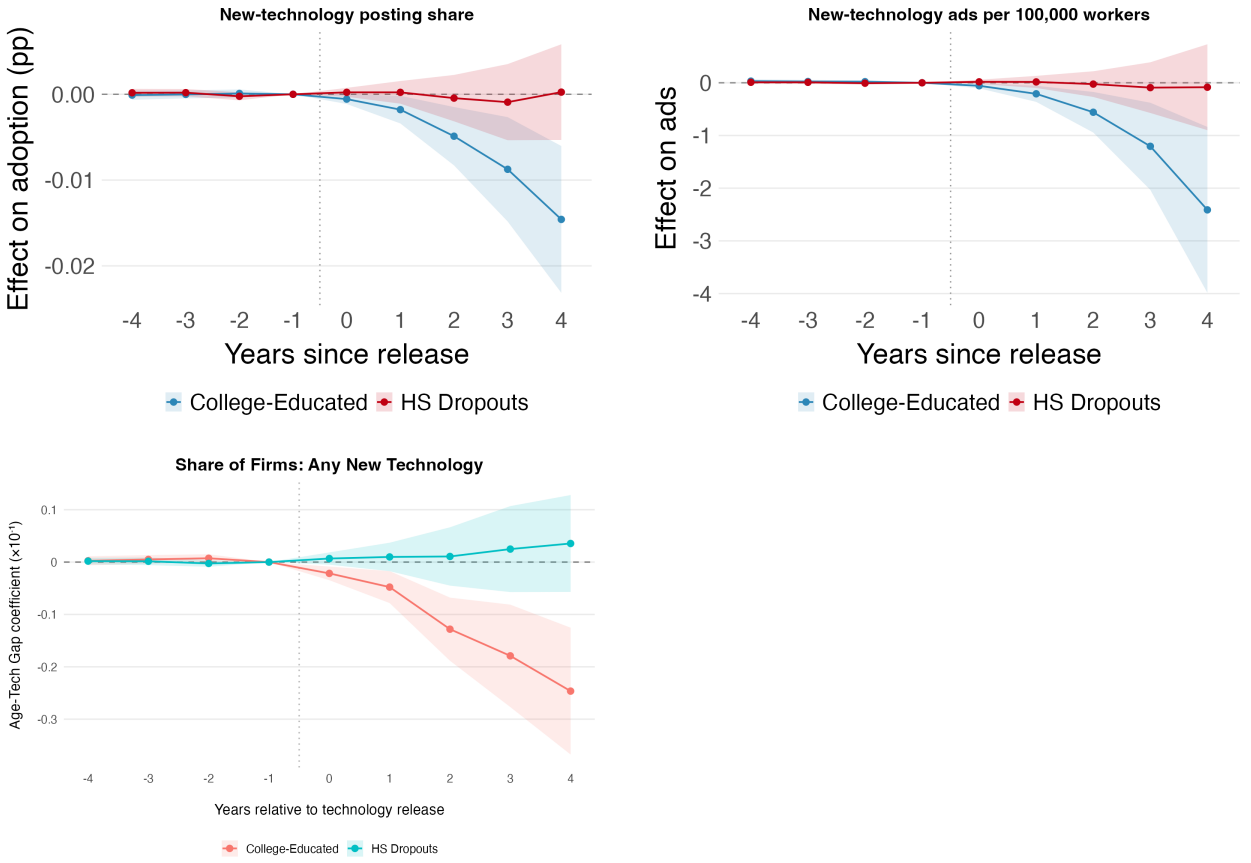


Figure 20: IV event-time estimates by education margin: college-educated versus less-than-high-school aging

Figure 21 reports a robustness check that augments the preferred state-by-year specification, beyond the baseline industry and college-share controls, with the local share of managers age 50 and older in 2011 interacted with event time. The manager-age coefficients are indistinguishable from 0 across all outcomes, while the estimated coefficients on the college-educated old-worker share remain close to the baseline results.

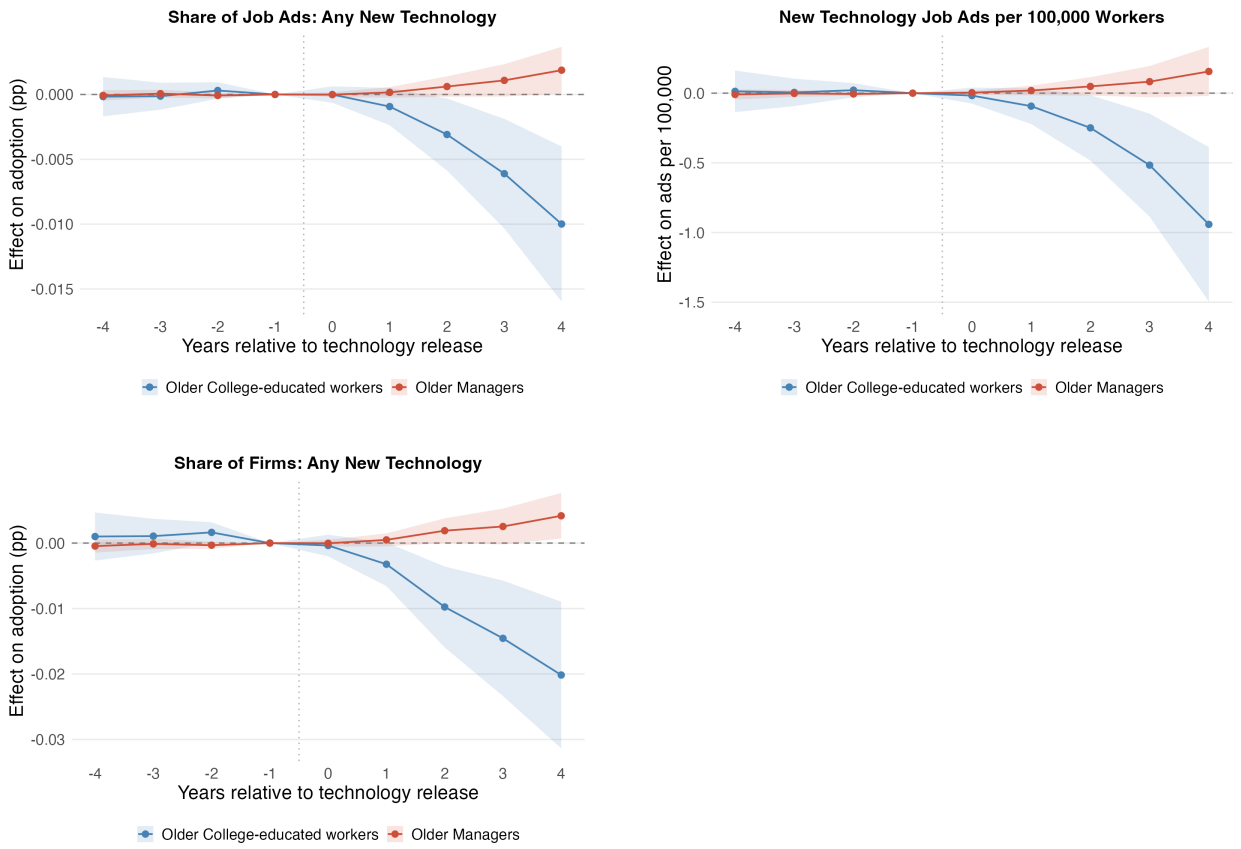


Figure 21: IV event-time estimates additionally controlling for manager age structure

H Event-Time First-Stage Diagnostics

Table 10 reports equation-specific effective first-stage F-statistics following Montiel Olea and Pflueger (2013). These statistics describe instrument strength in each first-stage equation.

Table 10: Event-time-specific effective first-stage statistics

Event time	Effective F
-4	104.6
-3	111.3
-2	115.1
0	118.1
1	117.7
2	115.8
3	110.9
4	102.9

Notes: The sample contains 84,455 observations in 665 commuting zones. The estimation sample is restricted to $-4 \leq \tau \leq 4$. All eight excluded event-time interactions enter each first stage; the reference period is $t = -1$. Controls and fixed effects match the preferred stacked IV specification. Let Z denote the excluded instruments residualized against these controls and fixed effects, $\widehat{\pi}_k$ the first-stage coefficients for endogenous interaction k, and $\widehat{u}_{ck}$ the corresponding residuals in commuting zone c. We report $F_{\text{eff},k} = \widehat{\pi}_k' Q \widehat{\pi}_k / \operatorname{tr}(Q^{-1}\widehat{S}_k)$, where $Q = Z'Z$ and $\widehat{S}_k = \sum_c Z_c' \widehat{u}_{ck} \widehat{u}_{ck}' Z_c$. Covariances are clustered by commuting zone, without finite-sample corrections.

I Rotemberg Weight Decomposition

Following Goldsmith-Pinkham et al. (2020), we assess which source cohorts actually identify the cohort-survival IV by decomposing the college-aging instrument into cohort-level Rotemberg weights. Figure 22 plots the absolute cohort weights ordered by source age. The four largest source cohorts account for about 59 percent of total absolute weight, and the six largest account for about 80 percent, so the identifying variation is quite concentrated. Lagged shares of older workers feed directly into the future old-worker share, while lagged shares of younger workers reduce it through the denominator.

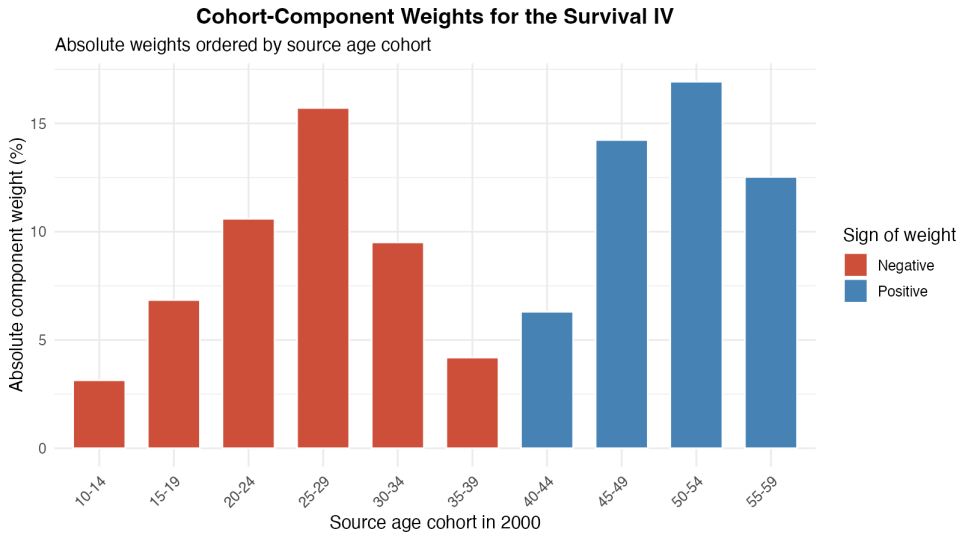


Figure 22: Absolute Rotemberg weights for the college-aging IV

J External Calibration Profiles

Figure 23 compares the controlled PIAAC age profile with the data and shows the implied learning-efficiency mapping. The score regression controls for sex, education, and survey wave. The estimated decline is 0.0171 standard deviations per year. The right panel translate this into learning profile under our current mapping assumption.

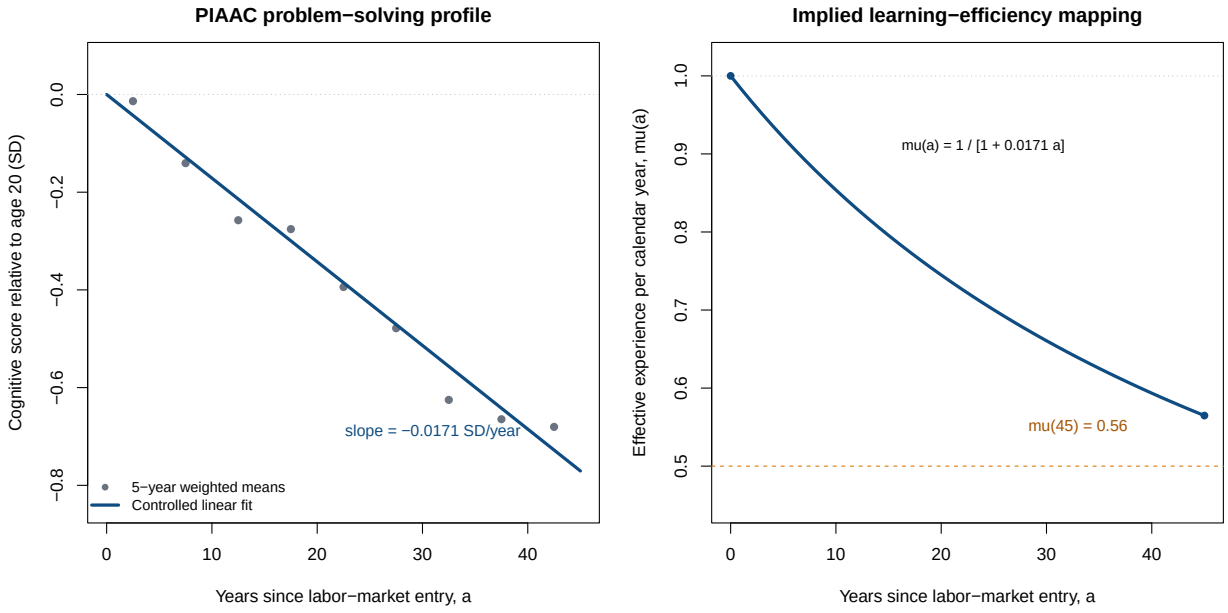


Figure 23: PIAAC age profile and implied learning efficiency

Figure 24 shows the original ACS experience–earnings fit used to obtain $K_{\max} = 0.883$ and $\rho_K = 0.131$. The sample comprises college-educated computer workers working full-time and full-year in the 2011 ACS. The age profile controls for commuting zone, occupation, education, sex, and race.

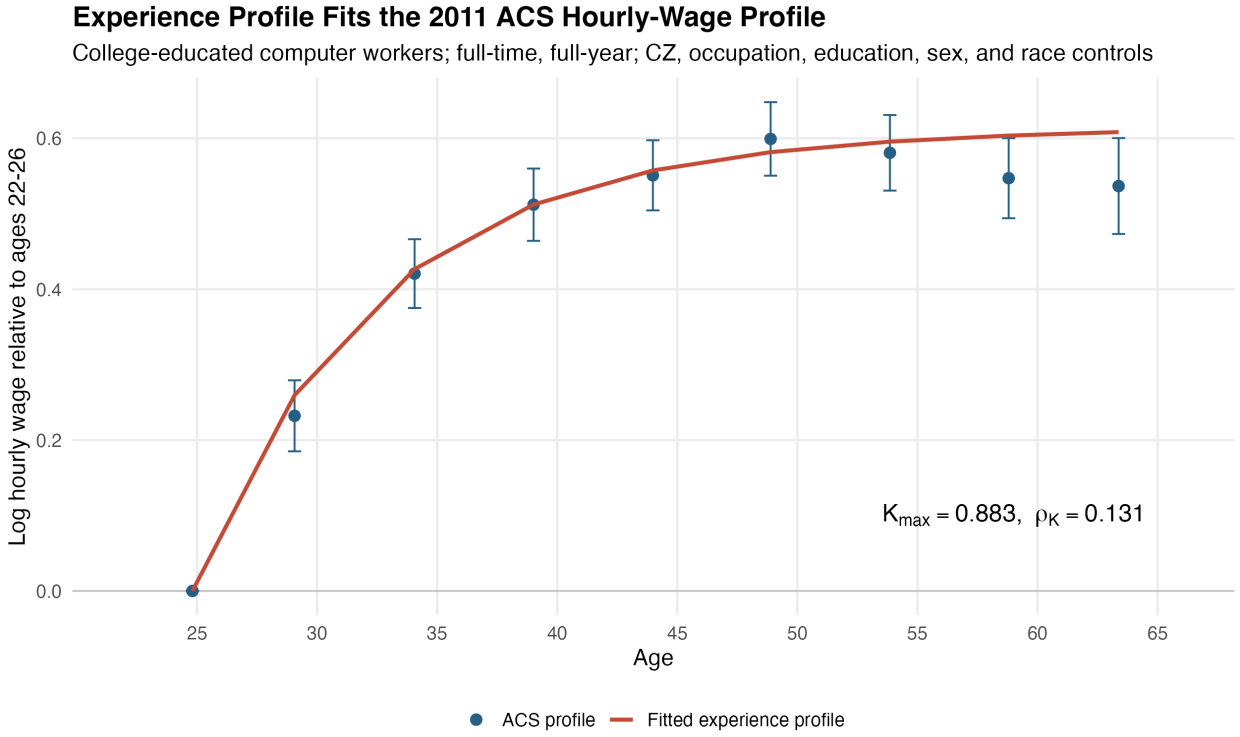


Figure 24: ACS age–earnings profile and the original experience-profile fit

Notes: Points are coefficients for five-year age groups, relative to ages 22–26; bars show 95% confidence intervals.

K Model Aggregation and Numerical Solution

K.1 Cohort Aggregation

Calendar time t measures years since the frontier arrives, while s measures years since a worker's technology choice. For an initial worker, these coincide. A worker entering at date $u > 0$, however, faces wages w_{u+s}. Define discounted earnings at decision date u by

$$\begin{aligned} B_I(a;u) &= \int_0^{T-a} e^{-rs} w_{u+s} K\Big(X_I(a,s)\Big)\, ds, \\ B_F(a;u,\phi) &= \int_0^{T-a} e^{-rs} w_{u+s} K\Big(X_F(a,s;\phi)\Big)\, ds. \end{aligned} \tag{15}$$

The aptitude cutoff at that date is

$$z^*(a;u,\phi) = \frac{B_I(a;u) + c(a,\phi)}{\alpha B_F(a;u,\phi)}. \tag{16}$$

At $u = 0$, this is the cutoff in equation (11). Future entrants have to make the decision at entry. Once chosen, the decision is irreversible.

For computation, we represent ages and entry cohorts at annual intervals. Let $n_0(a)$ denote the initial mass of workers aged a, with $\sum_{a=0}^{T-1} n_0(a) = 1$, and let n_u be the mass entering at date $u \geq 1$. Let $G(z)$ be the aptitude distribution, common across ages and entry cohorts. Define

$$q_I(x) = G(x), \qquad q_F(x) = \int_x^\infty z\, dG(z).$$

The first term is the fraction choosing the incumbent technology. The second is frontier aptitude summed over adopters per unit of cohort mass; it incorporates both the fraction adopting and their aptitude. With continuous aptitude, ties at the cutoff have zero probability.

Effective labor under the incumbent technology is

$$\begin{aligned} L_{I,t} = &\sum_{a=0}^{T-t-1} n_0(a)\, q_I\Big(z^*(a;0,\phi)\Big) K\Big(X_I(a,t)\Big) \\ &+ \sum_{u=\max\{1,t-T+1\}}^{t} n_u\, q_I\Big(z^*(0;u,\phi)\Big) K\Big(X_I(0,t-u)\Big), \end{aligned} \tag{17}$$

and effective labor under the frontier technology is

$$\begin{aligned} L_{F,t} = \alpha &\sum_{a=0}^{T-t-1} n_0(a)\, q_F\Big(z^*(a;0,\phi)\Big) K\Big(X_F(a,t;\phi)\Big) \\ &+ \alpha \sum_{u=\max\{1,t-T+1\}}^{t} n_u\, q_F\Big(z^*(0;u,\phi)\Big) K\Big(X_F(0,t-u;\phi)\Big). \end{aligned} \tag{18}$$

In each equation, the first sum covers initial workers who have not retired, and the second covers subsequent entrants still working at date t. Empty sums are zero. The initial-worker terms therefore disappear once $t \geq T$.

K.2 Numerical Solution

The default numerical horizon represents wage paths at annual dates from $t = 0$ through $t = 150$. Within each annual interval, wages are interpolated linearly, and discounted life-

time earnings are evaluated using four-point Gauss–Legendre quadrature. To evaluate the earnings of workers whose careers extend beyond the final date, we extend log wages at the rate $(\theta - 1)\log(1 + g)$, where g is the annual growth rate of entering cohorts. This is the wage growth rate implied by equation (5) when effective labor grows at rate g.

For each parameter vector, we initialize log wages at zero. Given a wage path, we calculate adoption cutoffs for initial workers and future entrants using equation (16), aggregate effective labor using equations (17)–(18), and obtain an implied wage path from equation (5). Let $\ell_{c,t}^{(k)}$ be the current log wage in commuting zone c, and let $\widehat{\ell}_{c,t}^{(k)}$ be the log wage implied by the resulting labor supply. We update wages according to

$$\ell_{c,t}^{(k+1)} = \ell_{c,t}^{(k)} + 0.45\left(\widehat{\ell}_{c,t}^{(k)} - \ell_{c,t}^{(k)}\right).$$

The default stopping criterion is

$$\max_{c,t}\left|\widehat{\ell}_{c,t}^{(k)} - \ell_{c,t}^{(k)}\right| < 2 \times 10^{-10}.$$

The solver allows at most 600 iterations and reports a failure if this criterion is not met.

L Alternative Calibration Without Experience Loss

This appendix reports the alternative calibration summarized at the end of Section 7.2.

Table 11 reports the results. Column (2) reproduces the baseline calibration from Table 3. Column (3) imposes no experience loss while holding learning efficiency at the PIAAC-based value $\lambda_\chi = 0.0171$. This specification cannot match the targets: the frontier advantage α is driven to its lower bound of one, the cross-market age coefficient is essentially zero (+0.001 against a target of −0.230), and the frontier earnings premium is overstated by roughly 60%. Without experience loss, and with learning efficiency declining only as fast as measured problem-solving performance, the model generates almost no cross-market relationship between workforce age and adoption.

Column (4) instead lets λ_χ adjust. The resulting parameter implies that a worker aged 65 is only 14.6% as effective in learning as an entrant worker, against 56.5% under the PIAAC mapping. The restricted model also fares worse on the untargeted moment. The standard deviation of frontier residual log earnings falls to 0.099 from the baseline 0.251, compared with 0.232 in the data. Additionally, the aptitude dispersion is also heavily compressed, down to 0.032 from 0.436. This exercise supports our overall conclusion that experience loss, not cognitive decline, is the major driver of the age-related adoption gap.

Table 11: Calibration without experience loss

			No experience loss	
	Data (1)	Baseline (2)	λ_χ at PIAAC (3)	λ_χ free (4)
Learning efficiency				
Cognitive decline λ_χ	—	0.0171	0.0171	0.1300
Multiple of PIAAC value	—	1.00	1.00	7.59
Learning efficiency at age 65	—	0.565	0.565	0.146
Estimated parameters				
Frontier productivity advantage α	—	1.517	1.000†	1.102
Aptitude dispersion σ_ε	—	0.436	0.081	0.032
Technology difficulty ϕ	—	1.579	0.614	0.804
Targeted moments				
Frontier posting share (%)	24.94	24.94	27.14	24.94
Cross-market age coefficient	−0.230	−0.230	0.001	−0.230
Frontier log-earnings coefficient	0.0779	0.0779	0.1234	0.0779
Untargeted moment				
Frontier residual log-earnings SD	0.2323	0.2514	0.1286	0.0988

Notes: Column (2) is the baseline calibration of Table 3. Columns (3) and (4) remove the loss of incumbent experience upon adoption. Column (3) fixes λ_χ at its PIAAC-based value while column (4) searches over λ_χ and reports the value at which all three targeted moments are matched.

M LLM Index Construction

The construction uses OpenAI GPT-5.4 for the rating while the cross-comparison provider uses Anthropic Claude Sonnet 4.6. This appendix describes the prompting protocol used to construct the difficulty and adoption-benefit indices. The implementation uses the same template for every legacy–modern technology pair, replacing only the legacy and modern technology names.

M.1 Difficulty Index

For each technology pair, the model is asked to act as an expert technical instructor designing a workforce training plan. It is given the legacy technology, the modern technology, and a common learner profile: a developer with two years of professional experience in the legacy technology, strong computer-science fundamentals, debugging ability, and approximately ten focused learning hours per week while working.

The model is asked to rate two components on a zero-to-five scale. The first component is

the conceptual shift: whether the move mainly involves syntax and API changes, or instead requires a substantially different mental model, architecture, or debugging approach. The second component is tooling and environment burden: whether the move uses essentially the same development environment, or instead requires a new operational stack, deployment workflow, observability setup, or infrastructure ecosystem.

M.2 Adoption-Benefit Index

For each technology pair, the model is asked to act as an expert software engineer assessing the practical benefit of the modern technology relative to the legacy technology in typical real-world production use. The model is instructed to evaluate the use cases that most plausibly motivate adoption of the modern technology, not to reward popularity or trendiness by itself, and to ground ratings in concrete technical differences.

The model rates six dimensions on an eleven-point scale from -5 to 5, where negative values mean the modern technology is deemed worse than the legacy technology, zero means no meaningful benefit, and positive values mean the modern technology is better. The six dimensions are:

1. code simplicity: reductions in code complexity, verbosity, boilerplate, or implementation burden;
2. runtime efficiency: improvements in speed, throughput, latency, or resource use;
3. feasibility at production scale: whether the modern technology makes important production tasks feasible, robust, or scalable when the legacy technology would require major workarounds;
4. maintainability: improvements in long-run support, tooling, libraries, documentation, testing, debugging, upgrades, observability, and developer availability;
5. safety and reliability: reductions in bugs, crashes, data corruption, security vulnerabilities, race conditions, memory errors, or operational failures;
6. flexibility: the ability to define custom logic, modify workflows, compose components, adapt to nonstandard requirements, or avoid reliance on fixed pre-built solutions.

For each pair, the benefit index averages the six dimension scores and then standardizes the resulting pair-level value across the fifteen technology pairs. Higher values therefore indicate modern technologies that offer larger practical production advantages relative to their legacy counterparts.

References

Daron Acemoglu and Pascual Restrepo. Demographics and Automation. *The Review of Economic Studies*, 89(1):1–44, January 2022. ISSN 0034-6527. doi: 10.1093/restud/rdab031.

Daron Acemoglu, David Autor, Jonathon Hazell, and Pascual Restrepo. Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, 40(S1):S293–S340, April 2022. ISSN 0734-306X, 1537-5307. doi: 10.1086/718327.

Yunus Aksoy, Henrique S. Basso, Ron P. Smith, and Tobias Grasl. Demographic Structure and Macroeconomic Trends. *American Economic Journal: Macroeconomics*, 11(1):193–222, January 2019. ISSN 1945-7707, 1945-7715. doi: 10.1257/mac.20170114.

Daniele Angelini. Aging Population and Technology Adoption. 2023.

David Autor, David Dorn, and Gordon Hanson. When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men. *American Economic Review: Insights*, 1(2):161–178, September 2019. ISSN 2640-205X, 2640-2068. doi: 10.1257/aeri.20180010.

David H. Autor and David Dorn. The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103(5):1553–1597, August 2013. ISSN 0002-8282. doi: 10.1257/aer.103.5.1553.

Tania Babina, Anastassia Fedyk, Alex He, and James Hodson. Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151:103745, January 2024. ISSN 0304-405X. doi: 10.1016/j.jfineco.2023.103745.

Yoram Ben-Porath. The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy*, 75(4, Part 1):352–365, 1967. ISSN 0022-3808. doi: 10.1086/259291.

Nicholas Bloom, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun. The Diffusion of Disruptive Technologies. 2021.

Lukas Buchheim and Martin Watzinger. The Employment Effects of Countercyclical Public Investments. *American Economic Journal: Economic Policy*, 15(1):154–173, February 2023. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.20180323.

David M Byrne, Stephen D Oliner, and Daniel E Sichel. Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

V. V. Chari and Hugo Hopenhayn. Vintage Human Capital, Growth, and the Diffusion of New Technology. *Journal of Political Economy*, 99(6):1142–1165, December 1991. doi: 10.1086/261795.

Diego Comin and Bart Hobijn. An Exploration of Technology Diffusion. *American Economic Review*, 100(5):2031–2059, December 2010. ISSN 0002-8282. doi: 10.1257/aer.100.5.2031.

Zara Contractor and Bledi Taska. The Effect of Software Adoption on Skill Demand and Wage Inequality. 2024.

Nicolas Crouzet, Pulak Ghosh, Apoorv Gupta, and Filippo Mezzanotti. Demographics and Technology Diffusion: Evidence from Mobile Payments, March 2024.

David J Deming and Kadeem Noray. Earnings Dynamics, Changing Job Skills, and STEM Careers*. *The Quarterly Journal of Economics*, 135(4):1965–2005, November 2020. ISSN 0033-5533. doi: 10.1093/qje/qjaa021.

Kevin S. Dubina. Labor force and macroeconomic projections overview and highlights, 2023–33 : Monthly Labor Review : U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/mlr/2024/article/labor-force-and-macroeconomic-projections-overview-and-highlights-2023-33.htm, 2024.

Andrew D. Foster and Mark R. Rosenzweig. Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture. *Journal of Political Economy*, 103 (6):1176–1209, 1995. ISSN 0022-3808.

Elisabeth Fürstenau, Niklas Gohl, Peter Haan, and Felix Weinhardt. Working life and human capital investment: Causal evidence from a pension reform. *Labour Economics*, 84:102426, October 2023. ISSN 0927-5371. doi: 10.1016/j.labeco.2023.102426.

Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift. Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8):2586–2624, August 2020. ISSN 0002-8282. doi: 10.1257/aer.20181047.

Zvi Griliches. Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica*, 25(4):501, October 1957. ISSN 00129682. doi: 10.2307/1905380.

Brad Hershbein and Lisa B. Kahn. Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings. *American Economic Review*, 108(7):1737–1772, July 2018. ISSN 0002-8282. doi: 10.1257/aer.20161570.

Hugo Hopenhayn, Julian Neira, and Rish Singhania. From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share. *Econometrica*, 90(4):1879–1914, 2022. ISSN 1468-0262. doi: 10.3982/ECTA18012.

Benjamin F Jones. Age and Great Invention. *Review of Economics and Statistics*, 92(1): 1–14, February 2010. ISSN 0034-6535, 1530-9142. doi: 10.1162/rest.2009.11724.

James Liang, Hui Wang, and Edward P. Lazear. Demographics and Entrepreneurship. *Journal of Political Economy*, 126(S1):S140–S196, October 2018. ISSN 0022-3808, 1537-534X. doi: 10.1086/698750.

Nicole Maestas, Kathleen J. Mullen, and David Powell. The Effect of Population Aging on Economic Growth, the Labor Force, and Productivity. *American Economic Journal: Macroeconomics*, 15(2):306–332, April 2023. ISSN 1945-7707. doi: 10.1257/mac.20190196.

Edwin Mansfield. Technical Change and the Rate of Imitation. *Econometrica*, 29(4):741, October 1961. ISSN 00129682. doi: 10.2307/1911817.

Paul Mohnen. The Impact of the Retirement Slowdown on the US Youth Labor Market. *Journal of Labor Economics*, 43(1):203–246, January 2025. ISSN 0734-306X. doi: 10.1086/725874.

José Luis Montiel Olea and Carolin Pflueger. A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics*, 31(3):358–369, 2013. ISSN 0735-0015. doi: 10.1080/00401706.2013.806694.

Raymond Montizaan, Frank Cörvers, and Andries De Grip. The effects of pension rights and retirement age on training participation: Evidence from a natural experiment. *Labour Economics*, 17(1):240–247, January 2010. ISSN 0927-5371. doi: 10.1016/j.labeco.2009.10.004.

Taraneh Shahin, María Teresa Ballestar de las Heras, and Ismael Sanz. Robots in action. *The World Economy*, 47(7):2828–2854, July 2024. ISSN 0378-5920, 1467-9701. doi: 10.1111/twec.13562.

James H. Stock and Motohiro Yogo. Testing for Weak Instruments in Linear IV Regression. In Donald W. K. Andrews and James H. Stock, editors, *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, pages 80–108. Cambridge University Press, Cambridge, 2005. doi: 10.1017/CBO9780511614491.006.

Michelle Yin, Hoa Vu, and Claudia Persico. How (un)Stable Are LLM Occupational Exposure Scores? Evidence from Multi-Model Replication, April 2026.