# Technological Innovations and Worker Skills

Matthew Nibloe [*] Peter Nunez [†] Joern Onken [‡]
Yilun Allen Shen [§]

October 27, 2026

[Please click here for the latest version]

**Abstract**

This paper studies how workers respond to skill-specific technological innovations through investment in training. We construct a novel measure of technological exposure by combining patent texts with O*NET skill descriptors using natural language processing. Linking this measure to UK longitudinal worker data, we find that greater technological exposure increases hourly wages, particularly among highly educated workers and those in non-routine occupations. Exposure also raises participation in training, especially training for future employment opportunities, and increases voluntary employer mobility. We develop a model of multidimensional skill investment under uncertain future skill returns. Higher skill returns provide more resources to pay for training, and allowing training to respond amplifies the output gains from technological change. Greater uncertainty about future skill returns, such as due to the unpredictable labour-market effects of AI, encourages workers to acquire a more balanced set of skills, giving them more flexibility to respond as skill returns change. These findings highlight workers' proactive investment in skills and pursuit of new employment opportunities as important margins of adjustment to technological innovation.

[*]University College London, IFS, RFBerlin and UCL Stone Centre. Email: matthew.nibloe.19@ucl.ac.uk
[†]London School of Economics. Email: p.nunez1@lse.ac.uk
[‡]University College London. Email: joern.onken.20@ucl.ac.uk
[§]University College London. Email: allen.shen.16@ucl.ac.uk

# 1 Introduction

Transformative technologies are often characterised by the worker skills that are most affected. While modern robotics now proficiently replicate manual skills such as physical strength and control precision, advances in artificial intelligence (AI) have subjected deductive reasoning and similar cognitive skills to technological disruption. When confronted with such skill-specific technological changes, how do workers adjust their human capital in response?

Centred on a multidimensional view of human capital, we study how workers respond to technological changes by connecting skill-level technological exposure to their adjustment of skill portfolio through training. To this end, we construct a novel measure of skill-level technological change by jointly analysing patent information and skill descriptions. Leveraging natural language processing techniques, we embed textual data of patents from Google Patents and skill descriptors from O*NET onto a common high-dimensional vector space. This allows us to quantify their pairwise similarity and infer technological innovation at the individual skill level.

In line with Acemoglu and Autor (2011), we view skills as the productive capabilities embodied in workers and used to perform tasks. While an occupation's tasks may change over time, skill dimensions provide a relatively stable basis for characterising human capital. Our approach complements the task-based view (Acemoglu et al., 2024; Gathmann & Schoenberg, 2010) by examining how technology changes the value of these underlying capabilities and workers' incentives to develop them through training.

The parsimony of the measure allows straightforward inference of technological exposure at the occupation and worker level. We aggregate skill-level exposure using occupational skill requirements and link the resulting measure to longitudinal data from the UK Household Longitudinal Study. Our empirical analysis exploits differences in the evolution of technological exposure across occupations to examine workers' subsequent adjustment in wages, employment, training and employer mobility. Using local projections, we estimate impulse response functions that trace how these outcomes evolve following an increase in occupational technological exposure.

We find that greater occupational exposure to technological change is associated with higher earnings, primarily through hourly wages. For a one-standard-deviation increase in exposure, the estimated increase in real hourly pay reaches approximately 3.1% at a horizon of ten survey waves, roughly ten years. Hours show little sustained adjustment, and there is no clear increase in unemployment. Together, these findings suggest that technological change raises the marginal product of labour on average, consistent with net complementarity. The wage gains are unevenly shared, with larger gains among

workers initially in non-routine occupations and those with higher education, consistent with the literature on technological change and wage inequality (Autor et al., 2003; Katz & Murphy, 1992).

The training and mobility findings suggest that workers adapt to the increased productive value of labour by developing skills for future employment and voluntarily pursuing opportunities elsewhere. Over ten survey waves, a one-standard-deviation increase in exposure is associated with 18.5 additional training reports per 100 workers. The reported purposes for training indicate that this investment has a forward-looking character: Training for future jobs increases, while the estimated change in training for current-job skills is small and imprecise. Over eight waves, the increase in future-job training reaches 4.1 additional training reports per 100 workers. Employer changes also increase, particularly voluntary moves for promotion or for a better job, with an estimated 3.7 additional reports of voluntary employer changes per 100 workers over the same eight-wave horizon.

To understand why wages and training increase together, we develop a model of multidimensional human capital investment under uncertainty, building on Cavounidis and Lang (2020). Workers hold two skills and choose how much to invest into each, as well as how to use them at work via occupational choice. Training uses resources that could otherwise be consumed today, while its benefits depend on the uncertain future skill returns.

Higher skill returns raise output and thus provide more resources for both consumption and training. In our numerical exercises, these additional current resources account for most of the increase in training expenditure, while changes in future payoffs mainly redirect investment towards the skill whose return improves. The effect on the composition of a worker's skill portfolio, which we label 'specialisation', depends on workers' starting skills. Investing more in the skill whose return improves can make a worker who already specialises in it more specialised, while making a worker who initially holds more of the other skill more balanced.

These investments also affect how much output rises following technological change. In an illustrative numerical exercise, allowing training to respond to a permanent improvement in one skill's return increases the discounted cumulative output gain by about a quarter at a horizon of twenty model periods. This is relative to retaining the investment path workers would have followed without the return improvement. More convex training costs reduce this amplification, suggesting that making training less costly for workers can amplify the output response to technological shocks.

We also ask how uncertainty affects the skills workers choose to acquire. Holding expected return paths and realised return histories fixed, we compare investment decisions made under different levels of forecast uncertainty about aggregate skill returns. In response

to greater uncertainty, workers acquire more balanced skill portfolios in this exercise, with only a small change in total training in the baseline. The diversification response also occurs under risk neutrality. Holding both skills allows workers to make greater use of whichever skill turns out to pay more. This relates to the current debate about the effects of AI on labour markets and degree choice. Our model suggests that increased uncertainty about which skill will be valuable in the future incentivises workers to follow a "wait-and-see" approach and thus not specialise into any particular skills. We plan to investigate the macroeconomic consequences of this result in ongoing work.

## 1.1 Related Literature

Our paper relates to the literature on technological change, skill demand, and wages. Katz and Murphy (1992) show how changes in skill supply and demand shape relative wages, while Autor et al. (2003) explain how technology changes demand by substituting for some tasks and complementing others. Evidence on broadband and robots documents wage gains and losses that vary across workers and labour markets (Acemoglu & Restrepo, 2020; Akerman et al., 2015; Graetz & Michaels, 2018). We find that technological change raises hourly wages on average, with larger gains among more educated workers and those in non-routine occupations. We build on this evidence by asking how workers adjust their human capital investment in response to changing technological opportunities.

Closest to our empirical approach, Kogan et al. (2023) link patents to job tasks and show that technologies that expand occupational labour demand can nevertheless reduce incumbent workers' earnings. Gathmann et al. (2024) examine how AI and robotics reshape tasks and workers' careers. These studies highlight the importance of workers' existing capabilities for adjustment to technological change. We build on the patent-based approach of Kogan et al. (2023) by linking innovations to O*NET skill descriptions and aggregating exposure using occupations' skill requirements. We use this measure to examine investment in human capital alongside changes in wages and employer mobility.

Our central empirical contribution concerns training as a response to technological change. Bartel and Sicherman (1998) find that young workers receive more formal training in industries undergoing faster technological change, while Bartel et al. (2007) link technology adoption to changing skill requirements and workplace training. Greater technological exposure need not induce more investment: Heß et al. (2023) find lower training participation among workers in occupations more exposed to automation. We contribute evidence on how investment responds over time and on whether workers train for their current work or future jobs. We find that technological change increases training, including preparation for future employment. Alongside the increase in voluntary employer changes, this pattern is consistent with workers developing capabilities to pursue new

opportunities.

Our model combines human capital investment with a multidimensional view of worker skills. As in Ben-Porath (1967), skill acquisition trades current resources against future earnings. Lazear (2009) shows how skills that are individually transferable can become specific in combination because firms place different weights on them. Lise and Postel-Vinay (2020) study multidimensional skills in a model of on-the-job search, with skills accumulating through use and depreciating when unused. We retain the distinction between skill stocks and their application, while making investment in each skill an explicit choice. Workers therefore choose both how much human capital to accumulate and which skills to develop.

We build on Cavounidis and Lang (2020), who model investment in multiple skills and allow workers to choose jobs that place different weights on them. Job choice creates returns to specialisation because workers can place more weight on the skills they develop. Convex investment costs provide a countervailing incentive to spread a given amount of investment across skills. We use a two-skill version of this framework with CRRA preferences and persistent stochastic skill-specific returns. Whereas their technological-change exercises consider unanticipated productivity changes, our workers account for future return risk when choosing investment. We distinguish the training response to higher current resources from the response to changes in future payoffs. We also study how uncertainty about these payoffs affects both the amount of training and the skills workers acquire.

The uncertainty margin is closely related to human capital portfolio choice. Silos and Smith (2015) study how uncertainty about occupational fit shapes the balance between specialised and diversified skill portfolios. In their model, workers choose their skills before labour market entry and subsequently learn about occupational fit. We instead study repeated training decisions under changing skill-specific returns. This connects our model to dynamic portfolio choice (Merton, 1969), with skills accumulated through costly investment rather than traded as financial assets. Holding several skills (i.e., a more generalist skill portfolio) allows workers to quickly put more weight on whichever becomes relatively more valuable as returns change. Risk aversion provides an additional reason to diversify, but is not necessary for this response in our numerical experiments; workers have a skill diversification incentive even under risk neutrality.

The remainder of the paper is organised as follows. Section 2 describes the data and the construction of our technological exposure measure. Section 3 documents patterns in technological exposure across skills and over time. Section 4 presents the empirical strategy, and Section 5 reports the results for wages, training and employer mobility. Section 6 develops a model of multidimensional skill investment and job choice under

uncertainty. Section 7 concludes.

# 2 Data

## 2.1 Skill-level Technological Exposure

We connect two data sources to construct a novel measure of skill-level exposure to technological changes. The foundation for our technology measure is the universe of patents granted by the United States Patent and Trademark Office (USPTO) from 1976 to 2020. We webscrape the abstracts of approximately seven million patents from Google Patents and retain those deemed economically salient, defined as ranking in the upper decile of cohort-adjusted citation counts.[1] Each patent abstract is parsed and embedded in a high-dimensional vector space using the all-mpnet-base-v2 sentence-transformer model, resulting in our technology text corpus $\mathcal{P}$.

In parallel, we draw on the Occupational Information Network (O*NET), of which the Content Reference Model deconstructs human capital into 87 distinct skill dimensions.[2] Each skill descriptor is concisely defined. For example, the skill descriptor "memory" is defined as "abilities related to the recall of available information". This allows us to embed the skill descriptors in the same vector space as the patents, producing our skill text corpus $\mathcal{S}$.

For any given O*NET skill descriptor $s \in \mathcal{S}$, its exposure to patent $p \in \mathcal{P}$ is measured by the cosine similarity between their respective vector embeddings, $Emb_s$ and $Emb_p$ :

$$C_{p,s} = \frac{Emb_p \cdot Emb_s}{\|Emb_p\|\|Emb_s\|}$$

Annual skill-specific exposure scores are then formed as the mean of cosine similarities across all patents filed in a given year, which yields a continuous, time-varying measure of the intensity with which each skill is affected by frontier technological advances:

$$X_{s,t} = \frac{1}{|\mathcal{P}_t|} \sum_{p \in \mathcal{P}_t} C_{p,s}$$

To construct occupation-level measures of technological exposure, we aggregate the skill-specific exposure scores using the O*NET data. O*NET quantifies the importance of

[1] This follows the conventional practice of applying natural language processing on patents data, e.g. Kelly et al. (2021), Kogan et al. (2019), Prytkova et al. (2024), and Webb (2019).

[2] These are descriptors classified as "abilities" or "skills". While the Content Reference Model includes nearly 300 descriptors, many of them pertain to work context or experience requirements, which we do not view as integral parts of human capital.

each skill for almost one thousand occupations, $\mathcal{J}$, covering the entire US economy, which we take as representative. For example, "mathematical reasoning" is scored 5 out of 5 (maximum importance) for mathematicians and 1 (minimum importance) for actors. We rescale these scores to construct harmonised skill-to-occupation aggregation weights $\omega_{o,s}$ for each occupation.[3] Figure 15 shows the relative weights on broad skill categories for each ISCO-88 major occupation group.

Using these weights, each occupation $o \in \mathcal{J}$ is then assigned a time-varying exposure score, which is the weighted sum of the skill-specific exposure scores:

$$E_{o,t} = \sum_{s \in \mathcal{S}} \omega_{o,s} X_{s,t}$$

This procedure closely resembles Kogan et al. (2023), which extends the semantic-matching approach pioneered by Webb (2019) by allowing exposure to evolve over time. It is important to note that we utilise a static version of the O*NET skill-to-occupation weights. As a result, any variation in the occupation-level exposure score over time reflects only changes in the underlying technological innovation at the skill level, rather than shifts in the skill composition of occupations. Since updates to the O*NET weights are infrequent and irregular, this approach remains a credible strategy for isolating the effects of technological change.

## 2.2 UK Household Longitudinal Study (UKHLS)

We combine Waves 1–18 of the harmonised British Household Panel Survey (BHPS) with Waves 1–14 of Understanding Society, the UK Household Longitudinal Study (UKHLS) (University of Essex, Institute for Social and Economic Research, 2026).[4][5] These representative household panels follow individuals at approximately annual intervals and record training alongside earnings, hours, employment status and occupation. Understanding Society Wave 14 includes adult interviews from approximately 22,000 households. The surveys are widely used to study UK earnings and labour-market dynamics (Blundell et al., 2021; Jolivet & Postel-Vinay, 2025; Postel-Vinay & Sepahsalari, 2023).

The surveys' detailed training information is a central reason for using this data. They cover activities such as part-time college or university courses, evening classes and employer-provided training. Respondents report why they train, allowing us to distinguish training to improve skills in the current job from training to prepare for future jobs. To our knowl-

[3] We utilise crosswalks between different occupation classifications to construct weights in line with the ISCO-88 classification, used by UK Household Longitudinal Study (UKHLS).

[4] The included interviews span 1991–2024. Understanding Society waves have overlapping fieldwork periods and span more than one calendar year.

[5] We utilise the Standard End User Licence (EUL) version of the dataset.

edge, this is the only representative longitudinal survey in the UK to combine training information at this level of detail with a long panel of individual labour-market outcomes.

We measure earnings as usual gross monthly main-job employee pay, annualised, and hourly pay as annualised earnings divided by 52 times usual weekly hours, including overtime.[6] Unemployment is measured by the respondent's main activity at the interview.

Drawing on Postel-Vinay and Sepahsalari (2023), we summarise job-spell histories between successive interviews and merge them with the individual-wave panel. These histories record employment between interviews, including changes of employer and reported reasons for leaving a job. They also capture moves that comparing employment status at successive interviews would miss. Employer changes include moves with an intervening period of nonemployment. We classify changes attributed to promotion or a better job as voluntary, and those attributed to redundancy, dismissal or the end of a temporary job as involuntary. As workers can change employer more than once between interviews, we may observe both voluntary and involuntary moves for the same person over this period. The two groups also do not include all reported reasons for leaving a job.[7]

We assign occupational exposure in a baseline wave and follow respondents aged 16–64 in that wave. Sample periods and projection horizons vary across outcomes according to the availability of comparable information. Table 1 presents survey-weighted characteristics of employees and self-employed workers in Understanding Society Waves 2–12, for the sample overall and for workers below and above the wave-specific mean of occupational exposure. The descriptive sample contains 221,856 individual-wave observations on 46,794 workers.

# 3 Technological Exposure: Stylised Facts

This section documents three stylised facts of the skill-level exposure measure $X_{s,t}$ constructed in Section 2.1. The aggregate series displays persistent growth with little relation to the business cycle; exposure has diverged across broad skill categories, with a split that opens around 2000; and it has diverged just as much between descriptors within the same category, indicating the technological shock a worker faces depends on the granular composition of that worker's skill portfolio. Comparisons across skills use each series indexed

[6] Pay outcomes require observed employee pay at both endpoints; we impose no separate restriction on baseline self-employment. Pay is deflated using the Consumer Prices Index for the actual calendar year of interview and expressed in 2015 pounds.

[7] We set the employer-change indicator to missing if we cannot establish whether the worker changed employer or stayed with the same one. This is also the case in Understanding Society Wave 1, for which employer changes cannot be identified. To distinguish between voluntary and involuntary moves, we additionally need the reported reason for leaving the job. Since workers can also be promoted without changing employer, we only include a promotion among voluntary moves if we also observe an employer change.

Table 1: Characteristics of employed workers, Understanding Society Waves 2–12

| | Overall | Below-mean exposure | Above-mean exposure |
|---|---|---|---|
| *Worker characteristics* | | | |
| Age | 41.9 | 41.3 | 42.4 |
| | (12.3) | (13.0) | (11.5) |
| Female (%) | 48.7 | 45.6 | 51.7 |
| Blue-collar occupation (%) | 24.8 | 49.9 | 0.9 |
| Non-routine occupation (%) | 73.7 | 70.7 | 76.5 |
| Higher education (%) | 45.9 | 29.3 | 61.6 |
| Part-time among employees (%) | 22.3 | 29.1 | 16.4 |
| *Pay and working time* | | | |
| Real hourly pay (£, 2015 prices) | 13.87 | 10.88 | 16.44 |
| | (61.94) | (54.80) | (67.37) |
| Annualised usual earnings (£, 2015 prices) | 25,447 | 18,543 | 31,418 |
| | (18,338) | (12,275) | (20,492) |
| Usual weekly hours, including overtime | 37.2 | 35.3 | 38.9 |
| | (13.1) | (13.7) | (12.2) |
| *Training* | | | |
| Any training (%) | 31.8 | 30.3 | 33.2 |
| Training for current-job skills (%) | 21.5 | 19.6 | 23.3 |
| Training for future-job opportunities (%) | 8.2 | 7.4 | 8.9 |
| *Employer mobility* | | | |
| Any employer change (%) | 8.4 | 8.5 | 8.3 |
| Voluntary employer change (%) | 4.5 | 4.2 | 4.7 |
| Involuntary employer change (%) | 1.7 | 1.8 | 1.7 |
| Total individual-wave observations | 221,856 | 107,721 | 114,135 |
| Distinct workers | 46,794 | 28,045 | 25,203 |

*Notes:* Survey-weighted means; standard deviations in parentheses for continuous variables. Indicators are percentages. The sample comprises employees and self-employed respondents aged 16–64 in Waves 2–12, with observed occupational exposure and positive cross-sectional weights. Groups use the worker-weighted mean exposure in levels within each wave; ties enter below. Weights are pooled without rescaling waves. Rows use available information. Pay is usual gross main-job pay in 2015 pounds; hourly pay uses usual hours including overtime. Part-time denotes employees with fewer than 30 positive usual weekly hours. Blue collar comprises ISCO-88 major groups 6–9; non-routine comprises groups 1–3, 5–6, and 9. Higher education denotes a degree or other higher qualification. Employer-change rates use observed interview intervals, including intervening nonemployment; reason categories can overlap and use their own observed denominators. Counts are unweighted; workers can enter both exposure groups across waves.

to its own 1976 value, $X_{s,t}/X_{s,1976}$, and levels are used only to describe the aggregate.

## 3.1 Aggregate exposure

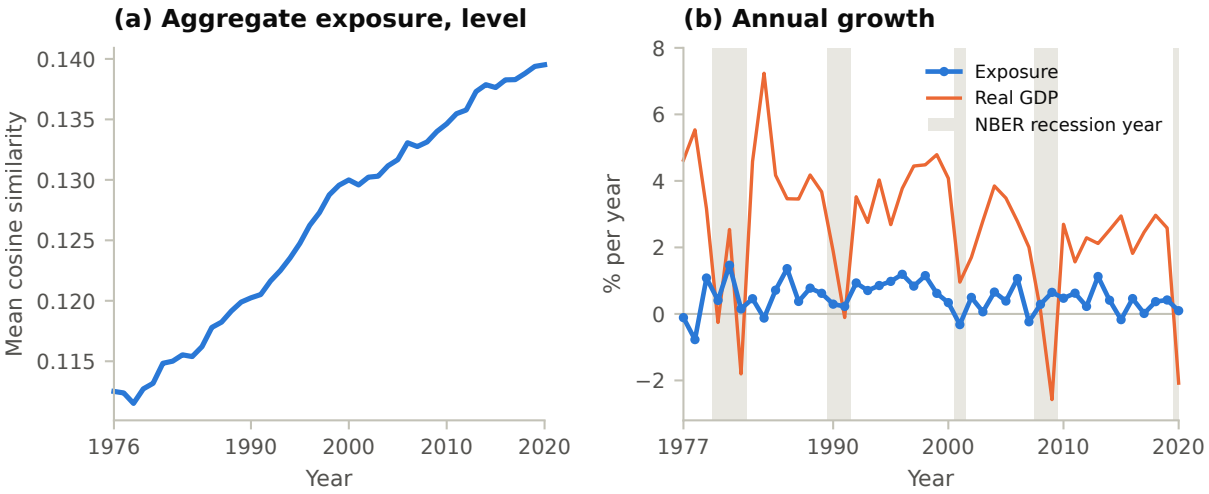



Figure 1: Aggregate technological exposure, 1976-2020

*Notes:* Panel (a) plots the mean of $X_{s,t}$ over the 87 O*NET descriptors (top-decile patents of each year, equally weighted). Panel (b) plots its annual growth rate (blue, with markers) together with the growth of US real GDP (orange), 1977-2020 (FRED). Shaded years contain at least one NBER recession month (FRED).

Averaged over all descriptors, exposure rose by 24% between 1976 and 2020, about 0.5% per year, with persistently positive growth (Figure 1a). Growth is also unrelated to the business cycle: its correlation with the growth of US real GDP is 0.08, and exposure kept growing through every recession in the sample except 2001 (Figure 1b). This is consistent with the view that technological breakthroughs are largely unpredictable and hence acyclical. It also indicates that the informative variation in the measure lies across skills rather than across years.

## 3.2 Divergence Across Skill Groups

Figure 2 plots the exposure of the 11 O*NET broad skill categories. Psychomotor and physical abilities recorded the largest increases, of roughly 50% over the sample period. These categories comprise bodily coordination, strength and movement, the capabilities that robotics and automation replicate, and the sample period encompasses the principal advances in these technologies. Social skills and complex problem solving recorded the smallest increases, of approximately 5%, all of which accrued before 2000. The most salient feature is the divergence that emerges thereafter. Whereas every category grew until 2000, only the embodied and equipment-related categories (psychomotor, physical and sensory abilities, technical skills and resource management) continued to do so; the categories associated with abstract reasoning, communication and interpersonal interaction

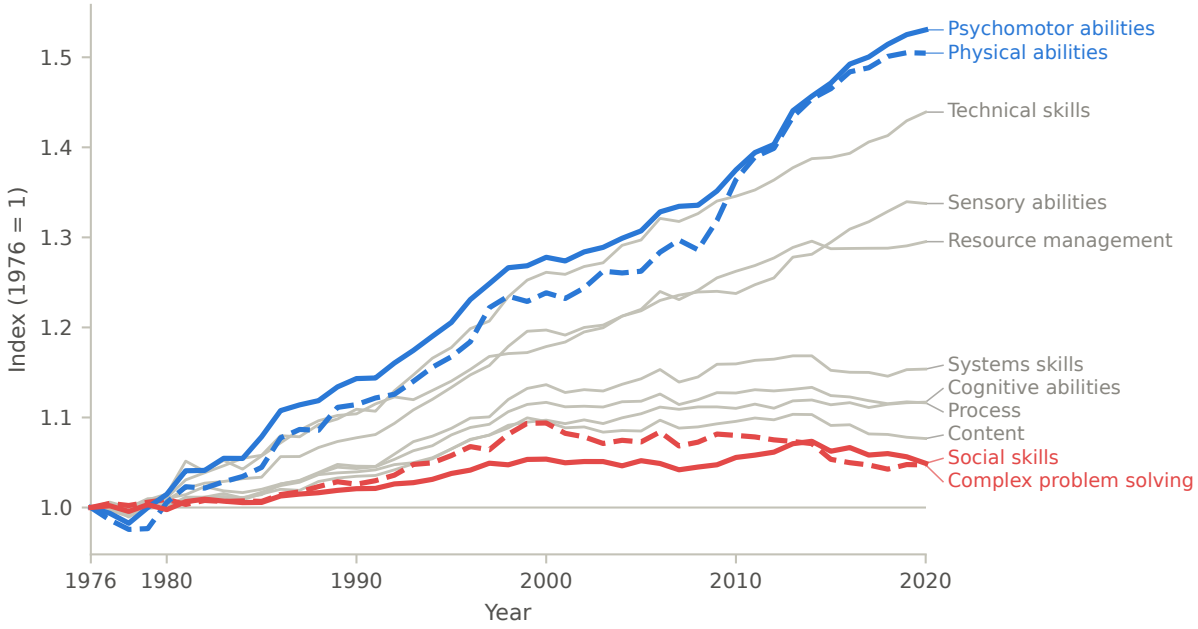



Figure 2: Exposure of broad skill categories, indexed to 1976

*Notes:* Each line is the mean of $X_{s,t}$ over the descriptors of one of the 11 broad O*NET categories, divided by its 1976 value. Blue lines denote the two fastest-growing categories and red lines the two slowest-growing; the remaining categories are shown in grey.

(cognitive abilities, systems and process skills, content, social skills and complex problem solving) stagnated or declined. Because occupations combine these categories in markedly different proportions (Figure 15), the divergence implies substantial heterogeneity in the shocks experienced by workers: exposure rose by roughly half for a portfolio concentrated in psychomotor and physical abilities, but remained essentially unchanged, and declined after 2000, for one concentrated in social and problem-solving skills.

## 3.3 Heterogeneity Within Skill Groups

A large share of the dispersion across the 87 descriptors arises within categories. Figure 3 documents this for three categories, highlighting the two fastest- and two slowest-growing descriptors of each. Within cognitive abilities, the 2020 index ranges from below one for fluency of ideas to above 1.5 for selective attention, a wider range than that spanned by the 11 category means; attention-related abilities grew fastest, whereas the two descriptors associated with creativity, fluency of ideas and originality, declined throughout the sample period. Within social skills, instructing rose persistently, by approximately one third, whereas negotiation and social perceptiveness were less exposed in 2020 than in 1976. The contrast separates instruction, an activity that technology increasingly performs, from activities that rely on interpreting and influencing other people. Within process skills, learning strategies and monitoring rose steadily, whereas active learning and critical thinking stagnated around 2000 and subsequently declined. In total, seven

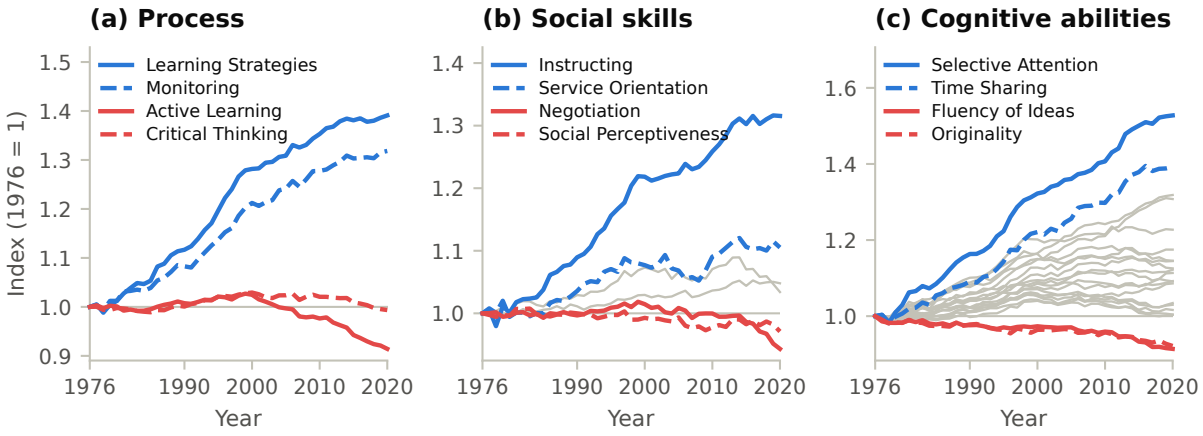



Figure 3: Exposure of granular descriptors in three categories, indexed to 1976

*Notes:* Each panel plots $X_{s,t}/X_{s,1976}$ for the descriptors of one broad category. Blue denotes the two fastest-growing descriptors of the category, red the two slowest-growing, and grey the remainder.

descriptors, drawn from four categories, were less exposed in 2020 than in 1976: fluency of ideas, originality, active learning, critical thinking, mathematics, negotiation and social perceptiveness. Their common theme is creativity, reasoning and learning, and interpersonal judgement.

Physical strength and manual dexterity, the skills replicated by robotics, are among the fastest-growing descriptors in our data, whereas reading comprehension and deductive reasoning, the skills most susceptible to advances in artificial intelligence, exhibit little growth by 2020. The patent-based measure therefore captures the diffusion of robotics and automation fully but that of artificial intelligence only partially, consistent with the concentration of the relevant patents at the end of the sample period. These within-category differences are also the reason the occupation-level measure aggregates descriptor-level scores with the descriptor-level weights $\omega_{o,s}$ of Section 2.1: aggregation to broad categories would eliminate much of the relevant variation.

# 4 Empirical Strategy

We study how workers' wages, training, and employer mobility respond to occupation-level technological change over time. Innovations recorded in patents may take time to affect the labour market, and workers may adjust gradually. We therefore use longitudinal data to trace responses at successive horizons using local projections (Jordà, 2005). Our approach of linking patent-based technological exposure to subsequent worker outcomes follows Kogan et al. (2023).

Our exposure measure links patents to skills and aggregates skill-level exposure to occupations using their skill weights. Each worker's exposure is assigned using their occupation

at baseline $t$. We then follow their outcomes over subsequent survey waves, including after changes of employer or occupation. Horizons are measured in survey waves, approximately one year apart.

## 4.1 Worker Outcomes

We examine real hourly pay alongside annualised usual real earnings, usual weekly hours, and unemployment to distinguish changes in pay from adjustment in working time and employment. For hourly pay, earnings, and hours, indexed by $k$, we measure the change in the log outcome between the wave before exposure and horizon $h$:

$$Y^k_{i,t,h} = \log y^k_{i,t+h} - \log y^k_{i,t-1}, \qquad h = 0, \ldots, 15. \tag{1}$$

The common reference period is $t-1$, so the response at each horizon describes the change accumulated since the wave before exposure. This allows us to trace the timing and persistence of wage responses.

Training is investment in skills. We are interested in the cumulative additional training activity following technological change, which captures investment undertaken over the response window. For count outcomes, let $D^k_{i,t}$ be a binary indicator for outcome $k$ at wave $t$, where $k$ denotes training, a training purpose, unemployment, or an employer change. We define the cumulative outcome as

$$Y^k_{i,t,h} = \sum_{j=1}^{h} D^k_{i,t+j}, \qquad h \geq 1. \tag{2}$$

For training, this counts the number of future waves in which the worker trains over the first $h$ waves after exposure. We examine any training over horizons one to ten. We also distinguish training intended to improve skills for the current job from training intended to prepare for future jobs, tracing each purpose over horizons one to eight.

For unemployment, $D^k_{i,t+j}$ indicates that unemployment is the respondent's main activity at the future interview. The cumulative outcome therefore measures the number of interviews at which the worker is unemployed, over horizons one to fifteen, among those employed or self-employed at baseline. For employer changes, the outcome counts future waves containing a change of employer over horizons one to ten. Employer changes include moves with intervening periods of nonemployment. We examine voluntary and involuntary changes separately to distinguish the reasons underlying mobility.[8]

We also examine how wage responses differ between workers in routine and non-routine

[8] Training purposes and reported reasons for employer changes can overlap; these categories are not additive decompositions.

occupations and between those with and without higher education.[9] Occupational type and educational attainment are measured at baseline, allowing us to compare responses across the groups initially exposed to technological change.

## 4.2 Specification

Let $E^{\text{std}}_{o,t}$ denote the occupational exposure score $E_{o,t}$ defined in Section 2.1, assigned to occupation $o$ in baseline wave $t$, residualised with respect to occupation and exposure-year fixed effects and standardised to unit variance across occupation–exposure-year cells. For each outcome and horizon, suppressing the outcome index $k$, our specification is

$$Y_{i,t,h} = \beta_h E^{\text{std}}_{o(i,t),t} + \alpha_i^{(h)} + \lambda^{(h)}_{c(i,t)} + \mu^{(h)}_{o(i,t)} + \varepsilon_{i,t,h}, \tag{3}$$

where $o(i,t)$ is individual $i$'s occupation at baseline and $c(i,t)$ is the calendar year of the baseline interview. The individual fixed effects $\alpha_i^{(h)}$ account for time-invariant differences across workers, the calendar-year fixed effects $\lambda^{(h)}_{c(i,t)}$ account for common changes across interview years, and the occupation fixed effects $\mu^{(h)}_{o(i,t)}$ account for permanent differences across baseline occupations.[10] Each horizon has its own coefficient and fixed effects.

The coefficient $\beta_h$ describes the response to a one-standard-deviation increase in residualised occupational exposure. For log outcomes, $100\beta_h$ gives the approximate percentage response from the wave before exposure to horizon $h$. For cumulative outcomes, $100\beta_h$ gives the additional number of waves with training, unemployment, or an employer change per 100 workers over the response window. The sequence of coefficients traces how these responses develop over time.

Identification relies on variation in occupational exposure over time, conditional on individual, calendar-year, and occupation fixed effects. We assume that, conditional on the included fixed effects, occupational technological exposure is exogenous to unobserved determinants of workers' outcomes at each horizon. We therefore interpret the responses as causal under this maintained assumption. Our motivation for treating exposure as plausibly exogenous is that individual workers do not explicitly incorporate annual patent flows into their labour-market decisions.

We interpret the coefficients as total dynamic responses to occupation-level technological change indexed by exposure at $t$. Exposure is persistent, and we allow subsequent exposure to evolve as part of the response of interest. The estimates therefore do not hold later exposure fixed to isolate a single year's increase. The specification also does not condition

[9] Higher education comprises a degree or other higher qualification, including higher-education diplomas and HNC/HND qualifications.

[10] Exposure years retain the year assigned to each survey wave; interview calendar years can differ within a wave.

on exposure history, so it does not separately identify the contribution of earlier exposure.

### 4.3 Estimation

We estimate equation (3) separately for each outcome and horizon, using longitudinal survey probability weights at $t+h$, which adjust for differential attrition. Wage changes require observed outcomes at $t-1$ and $t+h$, while cumulative outcomes require observations throughout the future window. Samples consequently vary by outcome and horizon, with longer horizons drawing on fewer observations and earlier baseline cohorts.

For the wage estimates, we further assume that, conditional on the included fixed effects, selection into the observed wage sample does not bias the estimated responses to exposure. The limited sustained responses of hours and unemployment reported in the Results section provide suggestive support for selection bias playing a limited role, consistent with technological change operating primarily through hourly wages.

Standard errors are clustered at the baseline occupation–exposure-year level, reflecting common exposure among workers in the same occupation and assigned exposure year. Formal comparisons of routine and non-routine wage responses use a joint specification with group-specific fixed effects. We report pointwise 95% confidence intervals.

## 5 Empirical Results

Greater occupational technological exposure raises hourly wages on average, with larger gains among workers in non-routine occupations and those with higher education. Workers also train more and make voluntary changes of employer. The increase in training for future jobs gives this adjustment a forward-looking character: workers invest in skills in preparation for employment opportunities they expect to pursue. Read together, these findings suggest that technology raises the marginal product of labour and workers adapt by developing their skills and changing where they use them. Throughout, responses correspond to a one-standard-deviation increase in occupational exposure after removing occupation and exposure-year fixed effects. Horizons are measured in survey waves, approximately one year apart. Appendix Tables B2–B5 report the estimates, standard errors, and sample sizes at selected horizons.

### 5.1 Wages and Employment

Greater exposure raises hourly wages, with gains that build over time. We examine these gains alongside earnings, hours, and unemployment to understand how technology affects the value of labour and workers' employment outcomes. Figure 4 reports the four responses. Hourly pay rises by approximately 1.0% at horizon five and 3.1% at horizon

ten. The gains persist through the end of the fifteen-wave window, when the estimate reaches 4.9%, with a 95% confidence interval of 1.4% to 8.3%.

Earnings gains emerge more gradually. Annualised usual earnings rise by approximately 4.5% at horizon twelve and 7.0% at horizon fifteen. The earlier response of hourly pay points to the wage margin as central to the earnings gains.[11] Hours fall modestly at intermediate horizons, consistent with an income effect from higher wages, but show little evidence of a sustained change.

Higher wages are not accompanied by a clear increase in unemployment. Panel (d) reports the cumulative number of future interviews at which a worker is unemployed. The response is initially negative and remains close to zero over the fifteen-wave window. At horizon fifteen, the estimate is 0.4 additional unemployment reports per 100 workers, with a 95% confidence interval of $-3.7$ to 4.4. Longer horizons draw on fewer complete histories and earlier baseline cohorts.

We interpret the increase in hourly wages, together with limited sustained adjustment in hours and unemployment, as evidence that technology primarily raises the marginal product of labour. Workers receive higher pay for their labour, with little sustained change in working time. This increase in the productive value of labour provides the starting point for our examination of how workers invest in skills and choose where to use them.

## 5.2 Heterogeneous Wage Effects

The average wage gain leaves open how the benefits of technological change are distributed across workers. The task-based literature emphasises differences between routine and non-routine work (Autor et al., 2003; Goos & Manning, 2007; Michaels et al., 2014). We examine this distinction alongside educational attainment to assess whether our wage responses display the patterns documented in that literature.

Wage gains are concentrated among workers initially in non-routine occupations. Panel (a) of Figure 5 shows an increasingly positive response for this group, while the routine-occupation profile is predominantly negative. At horizon ten, the estimates are approximately 5.8% and $-2.6$%, respectively. The gap is 8.3 percentage points, and a formal test rejects equal responses ($p < 0.001$). The positive average response therefore coexists with markedly different wage effects across occupational groups.

Educational attainment provides a second perspective on the distribution of wage gains.

[11] The pay and hours regressions use outcome-specific samples, so their coefficients need not add. On a common sample at horizon fifteen, the earnings response is 5.63%, equal to the hourly-pay response of 4.86% plus the hours response of 0.77%, up to rounding.

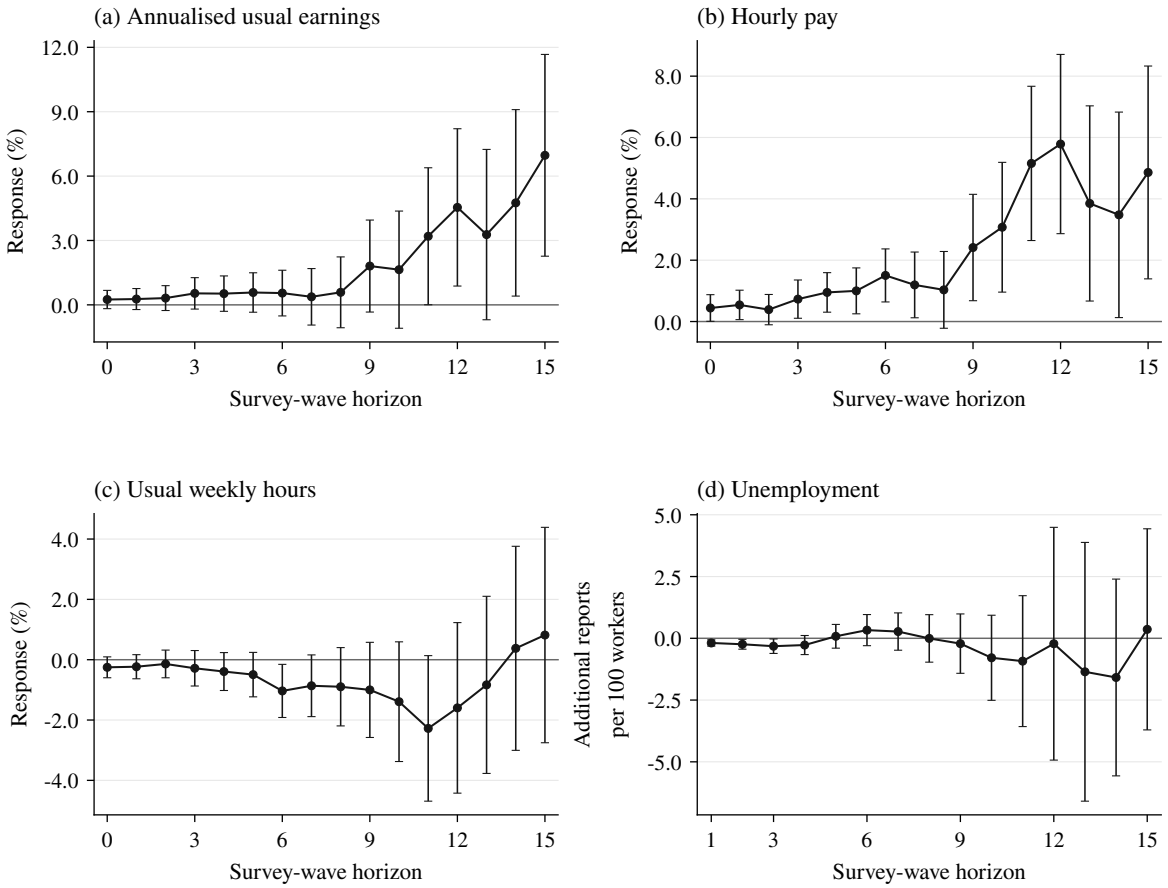



Figure 4: Wages and employment responses to occupational exposure

*Notes:* Panels (a)–(c) report changes in log outcomes from the wave before exposure to the indicated horizon, multiplied by 100; these are approximate percentage changes. Earnings are usual gross main-job pay annualised and deflated to 2015 prices. Hourly pay divides annualised earnings by 52 times usual weekly hours, including overtime. Panel (d) reports additional future interviews with unemployment per 100 workers employed or self-employed at baseline, measuring interview incidence rather than duration. Exposure is residualised on occupation and exposure-year fixed effects and standardised across occupation–exposure-year cells. Regressions include individual and baseline-occupation fixed effects and calendar-year fixed effects for the baseline interview, use longitudinal weights at the outcome horizon, and cluster standard errors by baseline occupation–exposure-year. Bars show pointwise 95% confidence intervals. Samples vary by outcome and horizon; future outcomes extend through the wave assigned to 2022.

In panel (b), workers with higher education experience gains that build over time. Their hourly-pay response rises from approximately 2.1% at horizon five to 6.0% at horizon ten and 11.8% at horizon fifteen. Although uncertainty increases at longer horizons, the estimated gains persist: the 95% confidence interval at horizon fifteen is 3.2% to 20.4%.

Workers without higher education experience smaller and less sustained gains. Their response remains close to zero through horizon eight, reaches approximately 1.6% at horizon ten, and is 1.1% at horizon fifteen. Positive gains appear at horizons eleven and twelve, but do not develop into the sustained increase estimated for workers with higher education. Taken together, the occupational and educational comparisons show that average wage gains are unevenly shared, with greater benefits for non-routine work and more educated workers. This pattern is in line with the literature on technological change and wage inequality (Katz & Murphy, 1992).

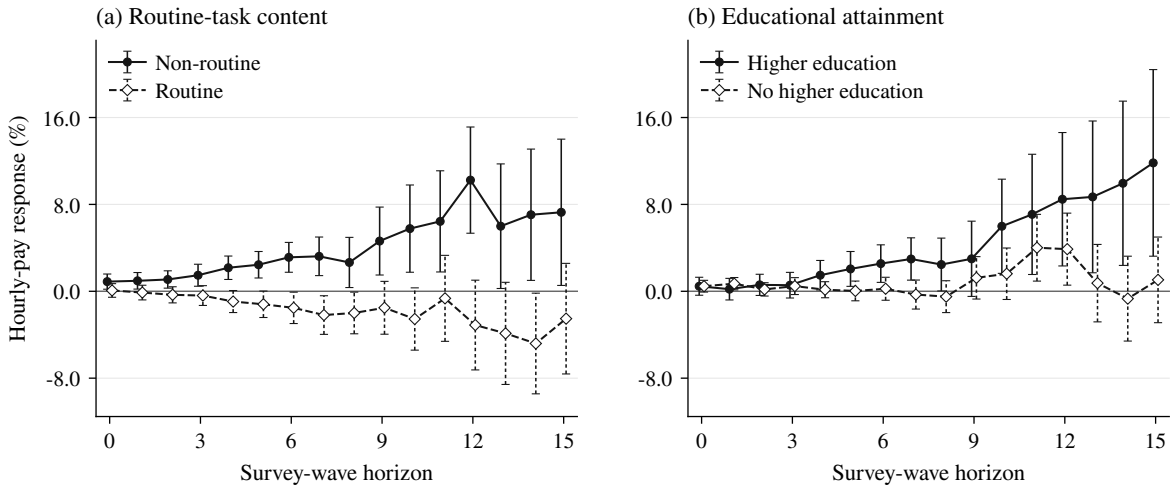



Figure 5: Hourly-pay responses by occupation and education

*Notes:* Each profile reports the change in log real hourly pay from the wave before exposure to the indicated horizon, multiplied by 100, for a one-standard-deviation increase in residualised exposure. Groups are assigned at baseline. Routine occupations are ISCO-88 major groups 4, 7, and 8; the remaining groups within 1–9 are classified as non-routine. Panel (b) compares workers with and without higher education. Missing group classifications are excluded from the relevant comparison. Solid lines with filled circles and dashed lines with open diamonds distinguish the groups. Estimation otherwise follows Figure 4. Bars show pointwise 95% confidence intervals. The formal routine-group comparison uses the corresponding joint model with group-specific fixed effects and the same occupation–exposure-year clusters.

The occupational classification is documented in Appendix Table B1, and Appendix Figure B1 reports the corresponding hourly-pay profiles separately for men and women.

## 5.3 Training

The wage results establish an increase in the productive value of labour on average. We now examine how workers respond by investing in their skills. Training builds human

capital and can prepare workers for the employment opportunities that technological change creates. Its reported purpose helps distinguish investment directed towards the worker's current job from preparation for future employment.

Training is common among employed workers. In Understanding Society Waves 2–12, 31.8% of workers in the descriptive sample report training in a given wave (Table 1). Training to improve current-job skills is more frequent than training for future-job opportunities, reported by 21.5% and 8.2% of workers, respectively. These rates describe the prevalence of training against which we examine workers' additional investment following an increase in exposure.

Greater exposure increases cumulative training participation. In Figure 6, the response of the number of future waves with reported training is positive throughout the window and increases with the horizon. After five waves, the response is 5.4 additional reports per 100 workers; after ten, it reaches 18.5, with a 95% confidence interval of 12.7 to 24.3. This is additional participation in an activity that many workers already undertake; the reported purpose of training helps explain how workers use that investment to adjust to technological change.

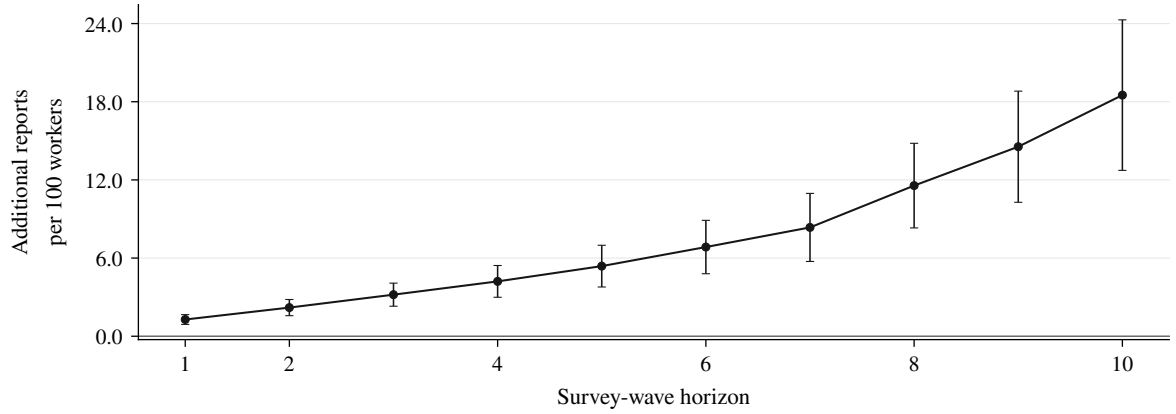



Figure 6: Training participation following occupational exposure

*Notes:* The outcome counts future waves with reported training over a fully observed response window, with at most one contribution per wave. Responses are additional positive training reports per 100 workers, rather than numbers of distinct courses. Any training pools BHPS and UKHLS. Exposure, fixed effects, longitudinal weights, clustering, and confidence intervals follow Figure 4.

The purpose comparison examines whether workers train to meet the demands of their present roles or prepare for future job opportunities. Figure 7 presents the two responses on the same vertical scale. Although training for current-job skills is more common in the descriptive sample, its response to exposure is small and imprecisely estimated. At horizon eight, the estimate is 2.1 additional reports per 100 workers, with a confidence interval spanning zero. Training for future-job opportunities rises more clearly at longer

horizons. By horizon eight, its response is 4.1 additional reports per 100 workers, with a 95% confidence interval of 1.7 to 6.6.

Promotion-related training provides a further forward-looking margin (Appendix Figure B2), complementing the voluntary employer changes examined below.

The moderate increase in reported training thus has a forward-looking character. Training for future jobs suggests that workers invest in skills for employment opportunities they anticipate beyond their present roles. In our interpretation, this investment helps prepare them for new jobs, making training one mechanism through which workers adjust their job choices to technological change. We next examine employer changes to study how workers act on those changing opportunities.

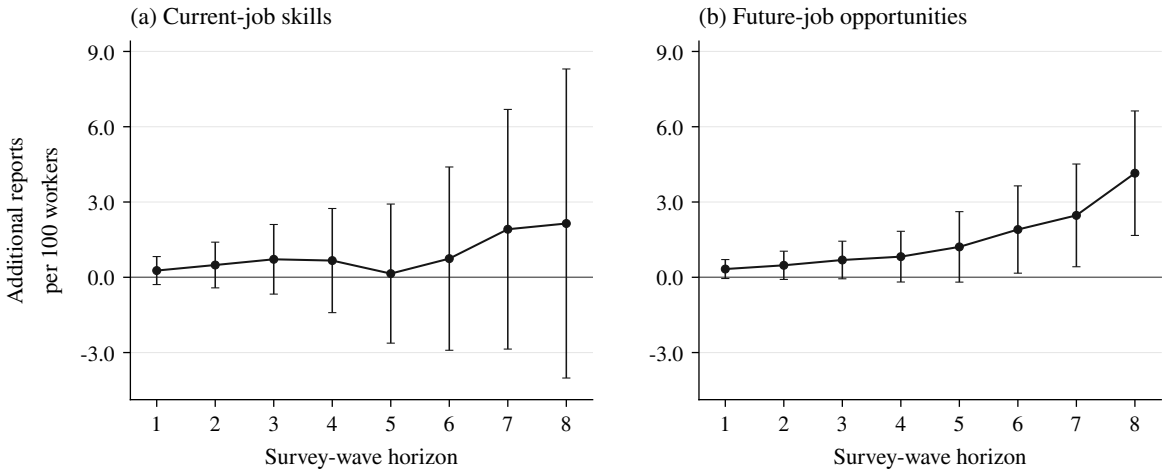



Figure 7: Training for current and future jobs

*Notes:* Both panels report additional positive reports per 100 workers and use the same vertical scale. Outcomes count future waves with a positive report of the indicated training purpose over fully observed windows. Purpose outcomes use UKHLS only. A wave may report both purposes, and observation requirements produce different samples across outcomes; the panels do not decompose the pooled any-training response. Estimation follows Figure 4; bars show pointwise 95% confidence intervals.

## 5.4 Job Switching

Training for future jobs prepares workers to pursue opportunities beyond their present roles. Employer changes provide a way to put those skills to use in a new job. The reported reasons for moving help distinguish voluntary moves to take advantage of new opportunities from moves associated with involuntary job endings. We study the cumulative number of future waves containing an identified employer change, allowing for repeated adjustment over the observation window.

Appendix Figure B3 places employer changes alongside broader measures of job and occupation mobility.

Occupational exposure increases employer mobility, with a more pronounced response at intermediate horizons. By horizon eight, panel (a) of Figure 8 shows 3.2 additional reports of an employer change per 100 workers, with a 95% confidence interval of 1.1 to 5.3. The estimate remains positive at horizon ten, although it is less precise.

The reported reasons for moving point towards voluntary adjustment. At horizon eight, the estimate for voluntary employer changes is 3.7 additional reports per 100 workers, compared with 1.0 for involuntary changes. The voluntary response remains positive at horizon ten, at 4.4 additional reports per 100 workers. Estimates for involuntary employer changes are generally smaller, with most confidence intervals including zero. Voluntary moves include promotions and departures for better jobs, so the more pronounced response in this category points to workers pursuing opportunities created by technological change.

Read together with the hourly-wage gains, the training and mobility responses describe workers adapting to the increased productive value of labour. Workers acquire skills for future employment and make voluntary moves to pursue new opportunities. In our interpretation, training helps prepare workers for those moves, linking human-capital accumulation to the choice of where to use it. The model formalises the interaction between technological productivity, skill investment, and job choice.

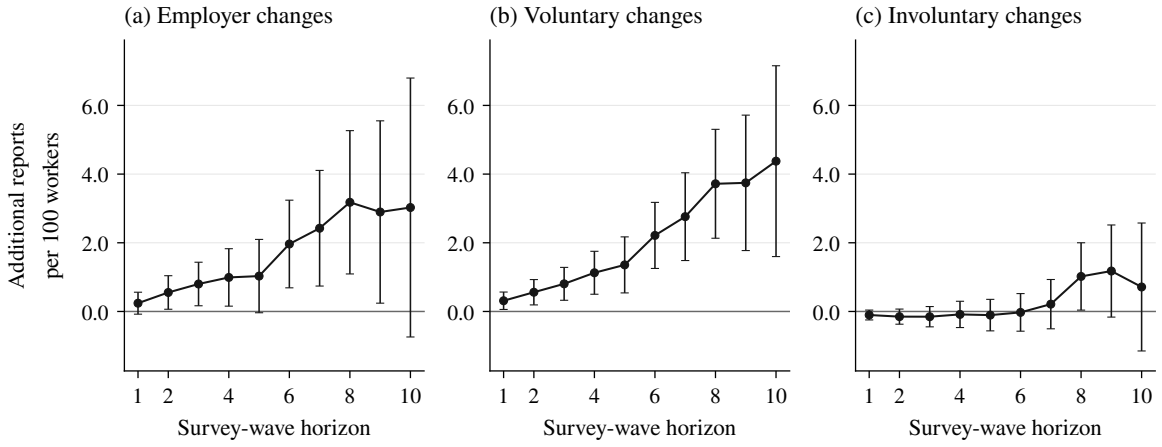



Figure 8: Employer changes and reported reasons for moving

*Notes:* The panels report additional positive reports per 100 workers for employer changes and their voluntary and involuntary categories, using the same vertical scale. Voluntary reasons include promotion or leaving for a better job; involuntary reasons include redundancy, dismissal, or the end of a temporary job. Employer changes include moves with intervening gaps as well as direct employment-to-employment transitions. Reasons can overlap within a wave and are not exhaustive; observation requirements also produce different samples across outcomes. The panels therefore do not form an additive decomposition of total employer changes. Exposure, fixed effects, longitudinal weights, clustering, and confidence intervals follow Figure 4.

# 6 A Model of Skill Investment under Uncertainty

Our empirical results show that workers train more and earn higher wages following an increase in technological exposure. We develop a model to understand why these responses occur together and how they depend on workers' existing skills. Higher skill returns provide more current resources to pay for training and change the future payoffs to acquiring skills. The local projections do not separate these two reasons for investing. The model allows us to do so, and to study how uncertainty affects both the amount of training and the skills workers develop.

We build on the multidimensional skill investment model of Cavounidis and Lang (2020). Workers hold two skills, choose how intensively to use them in production, and invest in each given uncertain future returns. With two skills, we can distinguish how much workers train from which skills they develop. They can build on an existing strength or acquire a skill they currently hold less of, with different implications for subsequent earnings and specialisation.

We first set out the model and study the response to higher returns. We then examine how uncertainty changes training and skill portfolios. Finally, we measure how much training adds to the output gains from a permanent return improvement.

## 6.1 Model Setup

**Skills and returns.** The economy contains a unit mass of infinitely lived workers, indexed by $n$, and two skills, indexed by $i \in \{1, 2\}$. Worker $n$ enters period $t$ with stocks $S_{n1,t}$ and $S_{n2,t}$. The effective return is $A_{ni,t} = B_{i,t}Z_{n,t}$. The component $B_{i,t}$ is common to all workers, while individual productivity $Z_{n,t}$ scales the returns to both skills of one worker. Both follow persistent processes specified in Appendix C.4. Workers observe current returns, know their transition probabilities and take them as given.

**Production and skill use.** We describe a job by the weights $J = (J_1, J_2)$ it places on the worker's skills. Suppressing worker and time indices, gross output is

$$Y = A_1 J_1 S_1 + A_2 J_2 S_2, \qquad J_1^\sigma + J_2^\sigma \leq 1, \quad J_i \geq 0, \quad \sigma > 1. \tag{4}$$

Workers choose these weights after observing returns. Changing the weights is costless, and they need not sum to one. A change can represent different tasks within a job or a move to another job. In the model, both are captured through the same choice, so the model does not separately determine employer-switching probabilities.

**Training and consumption.** Workers can increase either skill through investment $I_i \geq 0$. A fraction $\Delta_i \in (0,1)$ of each existing skill stock survives into the next period, giving

$$S_i' = \Delta_i S_i + I_i. \tag{5}$$

Training costs $K(I) = I_1^\rho + I_2^\rho$ units of the consumption good, with $\rho > 1$, leaving consumption $C = Y - K(I) > 0$. Spreading a given investment quantity $I_1 + I_2$ more evenly between both skills lowers its cost. Training quantity and expenditure can therefore move in different directions, and we distinguish both throughout this section. Training quantity is the closer counterpart to the observed training in our data, although our data only records the extensive margin of training.

We leave the division of training payments between workers and firms unspecified. Our model can thus also be interpreted as a worker-firm pair maximising the joint surplus. If firms pay a fixed share $\theta \in [0,1]$ of training costs, cash labour income is

$$w = Y - \theta K = (1-\theta)Y + \theta C. \tag{6}$$

For a given $\theta$, an increase in both output and consumption therefore raises labour income. The same is true for population averages when the payment share is the same across workers.[12] Hours are fixed and normalised to one, so labour income equals hourly wages in the model. This is consistent with our empirical finding that technological shocks increase hourly wages, not hours worked. Workers cannot save in financial assets or buy insurance, meaning their skills are their only savings device.

**The worker's problem.** Workers first choose how to use their skills and produce output. They then divide output between consumption and training, which adds to next period's skills. Since training costs do not depend on the job weights $J$, workers choose the weights to maximise current output. Substitution of the optimal weights $J$ into (4) gives

$$Y(S,A) = \left[(A_1 S_1)^q + (A_2 S_2)^q\right]^{1/q}, \qquad q \equiv \frac{\sigma}{\sigma - 1} > 1. \tag{7}$$

Appendix C.1 derives these $J$ weights. At fixed weights, output is linear in each skill. Allowing workers to put more weight on a stronger skill makes maximised output convex in skill stocks.

Workers discount future utility by $\beta \in (0,1)$ and have CRRA utility, with coefficient $\gamma \geq 0$. Utility is $u(C) = (C^{1-\gamma} - 1)/(1-\gamma)$ for $\gamma \neq 1$ and $u(C) = \log C$ for $\gamma = 1$. The

[12] This formulation is consistent with flexible wages and firms breaking even in each state and period, without insurance or intertemporal transfers. Let $K_w$ and $K_f$ be the worker's and firm's training payments. Zero profits give $w = Y - K_f$, and consumption is $C = w - K_w = Y - K$. The worker therefore bears the resource cost through lower consumption, regardless of who pays for the training (firm or worker).

investment problem is

$$V(S,B,Z) = \max_{I_1,I_2\geq 0} \big\{ u\big(Y(S,BZ) - K(I)\big) + \beta \mathbb{E}\left[V(S',B',Z') \mid B,Z\right] \big\},$$
$$S_i' = \Delta_i S_i + I_i, \qquad Y(S,BZ) - K(I) > 0. \tag{8}$$

Let $k_i = \rho I_i^{\rho-1}$ denote the marginal resource cost of investment in skill $i$. At a differentiable interior optimum, with positive investment in the relevant successor states, the Euler equation is

$$k_i = \beta \mathbb{E}_t \left[ \left( \frac{C}{C'} \right)^{\gamma} \left( Y'_{S_i} + \Delta_i k_i' \right) \right]. \tag{9}$$

An additional unit of investment raises future output, and the part that survives into later periods saves further training costs. The worker compares these benefits of skill investment with the cost of investing today, putting greater weight on benefits received in times of low consumption. Appendix C.1 derives the condition and the inequalities at investment corners.

**Measurement and parameterisation.** Reported training is the counterpart in the data to skill investment in our model. However, our data do not record the skill units acquired or their allocation, we only the extensive margin of training. The O*NET data describes occupational requirements rather than workers' actual own skill stocks. We therefore use the model to interpret the direction of the empirical responses and to study changes in workers' skill composition that our data cannot reveal. For example, training for a future job does not necessarily broaden a skill portfolio; it could also be used to further specialise. In other words, the effects of training on a worker's skill portfolio depend both on what the worker already knows (their current skill portfolio) and what they learn (the direction of training in the multidimensional skill space).

We measure individual skill specialisation by

$$\mathrm{sp}(S) = \frac{|S_2 - S_1|}{\sqrt{S_1^2 + S_2^2}}, \tag{10}$$

and set $\mathrm{sp}(0,0) = 0$. The index is zero at equal skills (the 45 degree line in our two-dimensional skill space) and one when only one skill is positive (perfect specialisation). Average (or aggregate) specialisation computes an unweighted average of the individual indices across workers, rather than applying the index to average skill stocks.

Our baseline parameterisation uses log utility, $\beta = 0.99$, $\sigma = 2$, $\rho = 2.5$ and skill retention rates $\Delta_1 = \Delta_2 = 0.95$. In our baseline stochastic return process, the common (skill-specific) and individual (skill-neutral productivity) log components each have persistence $\rho_A = 0.9$ and stationary standard deviation $0.95/\sqrt{2}$. This attributes half of the variance

of log effective returns to each component. We use these parameter values to illustrate the model's mechanisms; they are not calibrated to our data.

Our later dynamic temporary-return and uncertainty experiments use a different transition law for common returns, retaining the original levels and stationary probabilities. This lets us change uncertainty while holding expected level-return paths fixed. Appendix C.4 describes the parameterisation and solution method.

## 6.2 Adjustment to Higher Skill Returns

**Current income and future returns.** The same improvement in a skill's return, e.g. through a technology shock, affects workers differently depending on what skills they currently possess. From (7), the output elasticity and optimal relative skill use are

$$\begin{aligned} \frac{\partial \log Y}{\partial \log B_i} &= \frac{(B_i S_i)^q}{(B_1 S_1)^q + (B_2 S_2)^q} \equiv \omega_i(S, B), \\ \frac{J_1^*}{J_2^*} &= \left(\frac{B_1 S_1}{B_2 S_2}\right)^{1/(\sigma-1)}. \end{aligned} \tag{11}$$

An increase in the aggregate return of skill 1, $B_1$, raises output proportionately more for workers who rely more on skill 1. They can immediately put more weight on that skill, using stocks they already hold. Training affects later output, so its benefit depends on how long the return improvement lasts and how the acquired skills are subsequently used.

Higher returns can increase training by providing more current income or by making future investment payoffs more attractive. A two-period version of the model helps separate these effects. The worker has current resources $W > 0$, skills depreciate fully between periods, and there is no training in the final period. With log utility, optimal training expenditure and current consumption are

$$K^* = \frac{\beta W}{\rho + \beta}, \qquad C_0^* = \frac{\rho W}{\rho + \beta}. \tag{12}$$

Appendix C.2 derives these spending rules. Training expenditure is a fixed share of current resources. Hence, holding $W$ fixed, a change in future returns can change which skills the worker acquires without changing total spending. This holds for any admissible distribution of future skill returns. With full depreciation and log utility, spending also remains independent of future returns when training continues indefinitely, although the spending rule becomes $K^* = \beta W/\rho$.

With general CRRA utility, a common proportional increase in both future returns raises two-period expenditure when $\gamma < 1$, leaves it unchanged when $\gamma = 1$, and reduces it when $\gamma > 1$, at an interior optimum. Higher returns make investment more productive but also

allow workers to earn more from any given investment. When intertemporal substitution is low, the worker spends less on training. A common proportional increase in current resources and future returns instead raises expenditure for every $\gamma \geq 0$ at an interior optimum. Appendix C.2 derives these results.

In the full model, skills depreciate only partially, so future output depends on both the non-deprecated skills and new investment. The fixed-share spending rules derived above therefore no longer hold in general. The benchmark nevertheless shows that higher current income can finance more training even without an improvement in future returns. We therefore use the numerical model to separate the effect of having more resources today from the effect of better future payoffs to training, holding workers' initial skill stocks fixed.

**Training and wages after a return improvement.** We raise current $B_1$ by 10% for all workers, keeping initial skills, $B_2$ and individual productivity unchanged. The improvement is temporary but persistent: Its expected level advantage shrinks by 10% each period. At each horizon, we compare the affected population with the same workers without the initial increase, averaging over future common and individual histories. The reference population continues to adjust as well.

Under the reference process, a common return stays at its current level with probability 0.9 and is otherwise redrawn from its stationary distribution. We add the 10% higher value as a temporary state, avoiding an approximation by the next grid point. Appendix C.5 gives the construction.

Figure 9 shows that average investment in both skills rises on impact. Investment in skill 1 rises by 4.87%, while investment in skill 2 rises by 0.42%. The return improvement raises output by 5.12%, allowing workers to increase both training expenditure, by 6.79%, and consumption, by 4.25%. Hence, average hourly wages rise for any common fixed division of training payments between workers and firms. As the return improvement fades, the additional skills continue to raise output. After twenty periods, output is still 1.02% above the path without the improvement.

The model generates more training and higher wages, as in the empirical results. Comparing magnitudes between the empirical and model responses would require a mapping from patent-based exposure (our date) to skill returns (the model), and from reported training (extensive margin only) in the data to investment quantities in the model. Fixed hours and the absence of the employer-switching choice in the model also limit the empirical outcomes it can explain.

Furthermore, training is measured differently in the model and in the data. The empirical training IRF in Figure 6 reports cumulative training participation, while the model reports

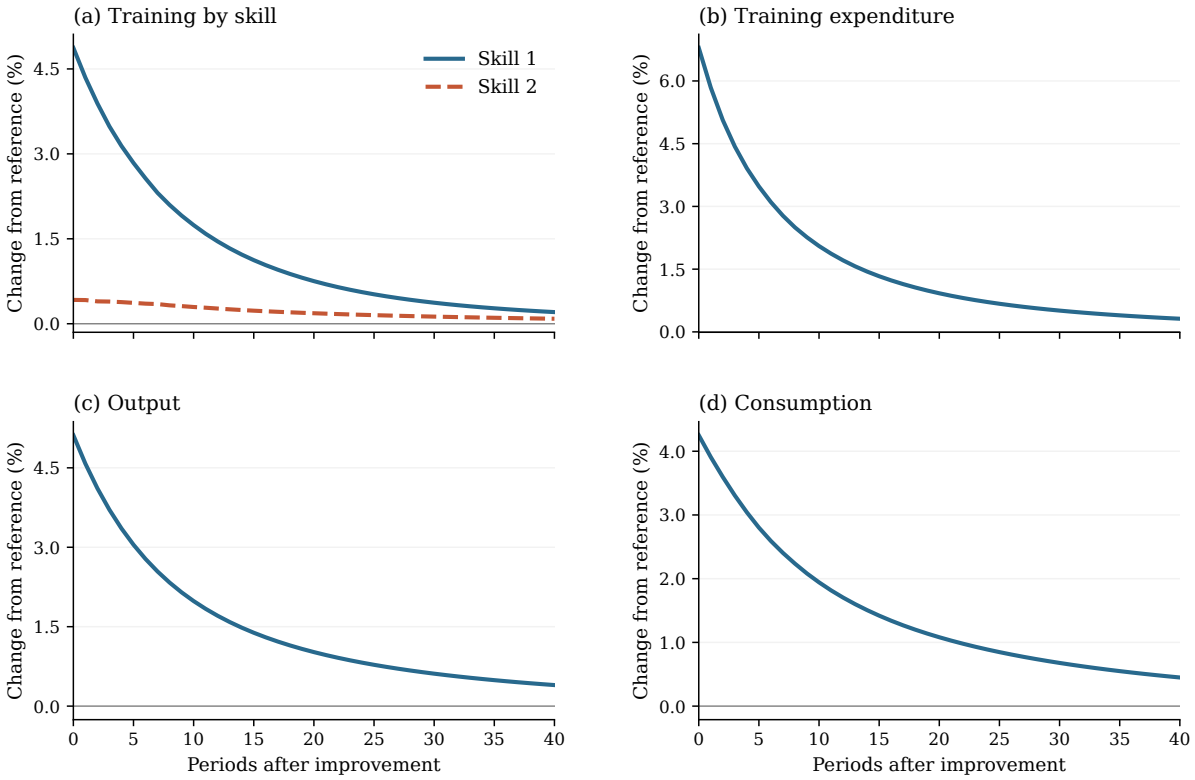



Figure 9: Training, output and consumption after a 10% skill-return improvement

*Notes:* Percentage differences in population means from the corresponding path without the return improvement. At period zero, $B_1$ rises by 10%, with initial skills, $B_2$ and individual productivity unchanged. Its expected level advantage shrinks by 10% each period. Panel (a) shows investment in each skill; panels (b)–(d) show expenditure $K = Y - C$, output $Y$ and consumption $C$. Means integrate over common and individual return histories.

investment per period. A positive but declining investment response in the model therefore continues to add to cumulative skill acquisition, but does not tell us how many workers participate in training or how often they train. Workers in the model observe improved returns and can train (adjust investment) immediately. Training participation in the data may respond more gradually if adopting technologies, learning about new opportunities or arranging training takes time. These delays are absent from the model. Changes in samples and weights across horizons also affect the empirical response.

**Current resources and future payoffs.** Table 2 separates the expenditure response into current resources and future payoffs. We use three portfolios with total skills $S_1+S_2 = 110$ and the middle one of 3 states for individual skill-neutral productivity. For each portfolio, we change current resources and the conditional future-return distribution separately and together, which yields four current investment problems with the same continuation-value function. Averaging each contribution across the two orders of introducing the changes makes the contributions add up to the total response. Appendix C.6 gives the definition.

Table 2: Current resources, future payoffs and training expenditure

| Initial skills $(S_1, S_2)$ | Current resources $\bar{R}_K/K_{00}$ (%) | Prospective payoffs $\bar{P}_K/K_{00}$ (%) | Total $\Delta K/K_{00}$ (%) |
|---|---|---|---|
| $(55, 55)$ | 7.01 | -0.23 | 6.78 |
| $(80, 30)$ | 12.34 | -0.42 | 11.92 |
| $(30, 80)$ | 1.77 | 0.38 | 2.15 |

*Notes:* Contributions as percentages of each portfolio's reference training expenditure after the same 10% increase in current $B_1$ as in Figure 9. Both $B_2$ and $Z$ start at their middle states. Within each row, initial skills are fixed across the four choices. The contributions average over the two orders of introducing the changes, as explain in equation (43) in the appendix. The future-payoff change affects the entire conditional return distribution, including its means and risk. Choices use the bilinear extension of the 600-point value function, with the same reference process and temporary return state as the dynamic experiment. Rounding can make displayed contributions differ from their total.

Most of the increase in expenditure comes from additional current resources. For the balanced portfolio, expenditure rises by 6.78%. Current resources contribute 7.01 percentage points of reference expenditure, while future payoffs subtract 0.23 points. Changing only the future-return distribution raises investment in skill 1 by 1.87% and reduces investment in skill 2 by 2.11%. When both channels operate, the extra current resources allow the worker to invest more in both skills.

The future-payoff contribution subtracts 0.42 percentage points at $(S_1, S_2) = (80, 30)$ but adds 0.38 points at $(30, 80)$. This is despite these two workers having equal total skills and specialisation; what differs is whether they hold more of the skill whose return improves (skill 1). The comparison changes the entire conditional return distribution; Appendix C.7 reports the numerical bounds.

**Differences across workers.** Figure 10 follows a balanced worker and two equally specialised workers, one with more of skill 1 and the other with more of skill 2. Their portfolios are grid counterparts of those in Table 2, with equal total skills of approximately 110.2 and identical initial returns and productivity. Each worker is compared with their own expected path without the skill-1-return improvement.

Output initially rises by 8.82% for the worker who is specialised in skill 1, 5.12% for the balanced worker and 1.29% for the worker who is specialised in skill 2. The two specialised workers have the same output at equal returns, so their different gains reflect which skill they rely on. These gains fade with the return advantage.

After twenty periods, specialisation is 0.16 percentage points higher for the worker initially concentrated in skill 1 and 0.80 points lower for the worker concentrated in skill 2, relative to their no-improvement paths. This is because investing towards skill 1 makes the first worker more specialised and the second worker more balanced. The balanced worker's specialisation rises by 0.14 points. Hence, both the degree and the direction of initial

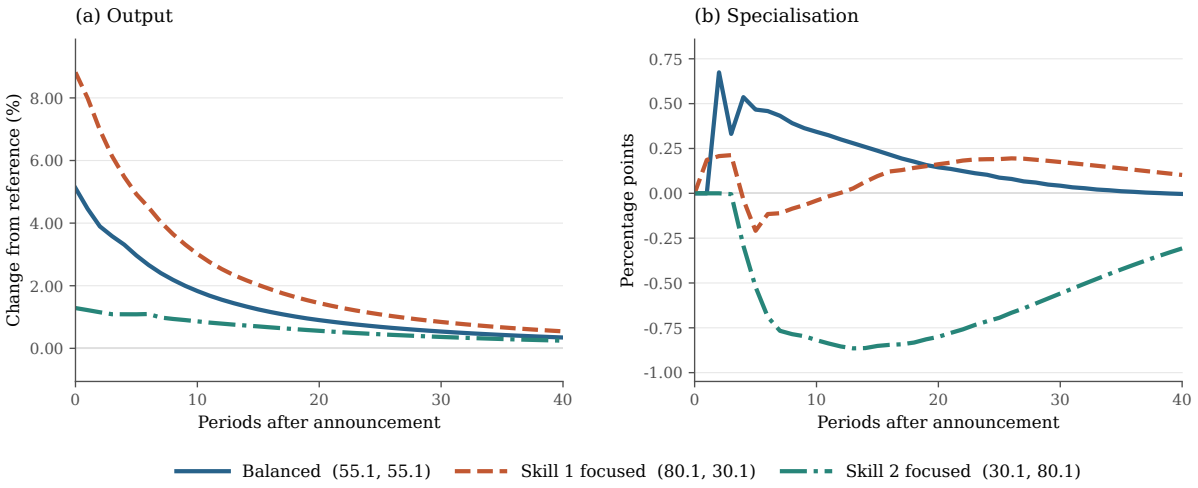



Figure 10: How initial skills affect the response to higher returns

*Notes:* Conditional expected responses to the temporary, persistent 10% increase in $B_1$ from Figure 9. Rounded initial portfolios are $(55.1, 55.1)$, $(80.1, 30.1)$ and $(30.1, 80.1)$, with equal total skills and middle-state $B_1$, $B_2$ and $Z$. Output is a percentage difference from each worker's own expected path without the increase. Specialisation is the expected individual index in (10), with changes in percentage points. Expectations use the same extended return process as the aggregate experiment. Appendix C.7 reports the comparison. Horizons are model periods.

specialisation matter.

The skills acquired through additional training continue to raise output as the return improvement fades. Appendix C.6 separates the output response into the direct effect of higher returns, changes in how workers use their skills and changes in their skill stocks. Accumulated skills account for about 33% of the remaining output gain after ten periods and 60% after twenty. As the direct return effect weakens, accumulated skills explain a larger share of the gain, even though the total output gain is shrinking. Figure C1 reports the accounting.

## 6.3 Uncertainty and Skill Diversification

Workers may change which skills they acquire when future returns become more uncertain, even when expected returns stay the same. In our numerical experiment, greater uncertainty about skill returns causes workers to adopt more more balanced skill portfolios. We isolate the change in workers' training decisions due to increased uncertainty by applying their optimal investment policies chosen under different skill return forecasts to the same realised return histories.

**Why uncertainty can encourage skill diversification.** Holding a more balanced (i.e., generalist) skill portfolio allows workers to put more weight on whichever skill turns out to pay more. This can be interpreted as workers adopting a "wait and see" approach. Risk aversion adds another reason for workers to diversify, as it makes income (which is

obtained through skills) more valuable in low-consumption states of the world. Note that changing the skill mix can only protect against relative-return risk, but cannot protect against individual productivity $Z$, as that affects both skills at the same time.

We can show the diversification response to increased uncertainty analytically in a two-period version of the model where workers keep their existing skills. Let current resources be $W > 0$ and retained skill stocks after depreciation be $d_i = \Delta_i S_i \geq 0$. The worker chooses $I_i \geq 0$ to maximise $\log[W - K(I)] + \beta \mathbb{E}[\log Y(d + I, A')]$ and consumes all output in the final period. Future returns have equal conditional means and follow the symmetric lognormal experiment in Appendix C.3. We increase the variance $v_R$ of log relative returns while keeping both expected level returns and independent common risk fixed.

**Proposition 1 (Uncertainty and skill diversification)** *In this two-period problem, let $\rho > q = \sigma/(\sigma - 1)$. At the unique optimum, define $x^*(v_R) = \log(S_1'^*/S_2'^*)$. For every finite $v_R \geq 0$,*

$$\operatorname{sgn}\left(\frac{dx^*}{dv_R}\right) = -\operatorname{sgn}(x^*). \tag{13}$$

*Hence, greater relative-return uncertainty makes an asymmetric chosen skill portfolio less specialised. The derivative at zero is a right derivative.*

The result allows inherited skills and positive initial uncertainty. If the chosen next-period stock of skill 1 exceeds that of skill 2, additional risk lowers investment in skill 1 and raises it in skill 2. Appendix C.3 proves the result. Equal expected returns are important: When expected returns differ, additional risk shifts investment away from the skill with the larger product of expected return and next-period stock, $\bar{A}_i S_i'$. This can be the skill with the smaller stock if its expected return is sufficiently high. The adjustment can therefore increase specialisation measured by skill stocks alone.

The proposition determines how uncertainty shifts investment between skills, but not whether total training expenditure rises or falls. With full depreciation and log-utility, (12) implies that expenditure remains fixed. Since training costs are convex, a more balanced investment allocation then allows workers to acquire more total skill units for the same expenditure. When workers retain existing skills, expenditure can also respond to uncertainty.

In the full model, workers continually train and face persistent differences in skill returns. The two-period result hence does not determine how uncertainty affects skill investment and skill holdings in the full model. We therefore examine these responses numerically, considering both investment decisions and the resulting skill distribution.

**The uncertainty experiment.** We permanently raise the conditional level variance of each common return by 20% in workers' forecasts, keeping individual-productivity risk

and expected level returns at every horizon unchanged. Appendix C.5 constructs the process. We solve the model under each forecast and evaluate both the normal-risk and higher-risk policies under the same realised reference process. The higher-risk investment policy is optimal under the higher-risk forecast, but is not necessarily optimal under the reference skill return process. Hence this exercise isolates how greater forecast uncertainty changes investment and subsequent skill accumulation, while holding the realised return histories fixed.

We compare training-cost curvatures of $\rho = 2.5$, 3 and 4. To make these comparisons, we normalise training costs as

$$K_\rho(I) = \kappa_\rho \left(I_1^\rho + I_2^\rho\right), \qquad \kappa_\rho = \frac{2.5}{\rho}\bar{I}^{\,2.5-\rho}, \qquad \bar{I} \simeq 2.81.$$

Here, $\bar{I}$ is investment in each skill at the baseline deterministic symmetric steady state. This normalisation keeps marginal cost fixed at the reference investment while changing its elasticity with respect to investment. It does not keep total expenditure or the eventual desired skill portfolio fixed. For example, at $\rho = 4$, expenditure at the reference investment is 37.5% lower than at $\rho = 2.5$.

All curvatures start from the same joint distribution of skills and individual productivity, conditional on the same common returns. Each higher-risk policy is compared with its own reference policy under the same realised histories. Subtracting the evolving reference removes adjustment that would occur anyway without the uncertainty increase.

At the baseline curvature, training quantity is 0.29% higher after one hundred periods, while expenditure is 0.46% lower. Workers divide investment more evenly, and the share of training quantity directed towards their currently smaller skill rises from 46.11% to 46.68%. The resulting saving outweighs the costs of greater mean training and dispersion in total quantities. Appendix C.6 gives the accounting.

These choices reduce mean specialisation by 1.47 percentage points at period one hundred and raise total skills by 0.29%. Output is nevertheless 0.30% lower: Skill quantities contribute 0.30 percentage points and portfolio composition contributes $-0.60$ points under the same realised returns. Most of the portfolio-composition-induced loss reflects weaker alignment with realised returns. Even when workers still hold more of the higher-return skill, a more balanced (less specialised) skill mix reduces its share and hence output per unit of skills. The rise in total skills is consistent with precautionary skill accumulation, but also reflects the lower cost of a more balanced bundle due to the convex training cost, so it does not separately identify that motive.

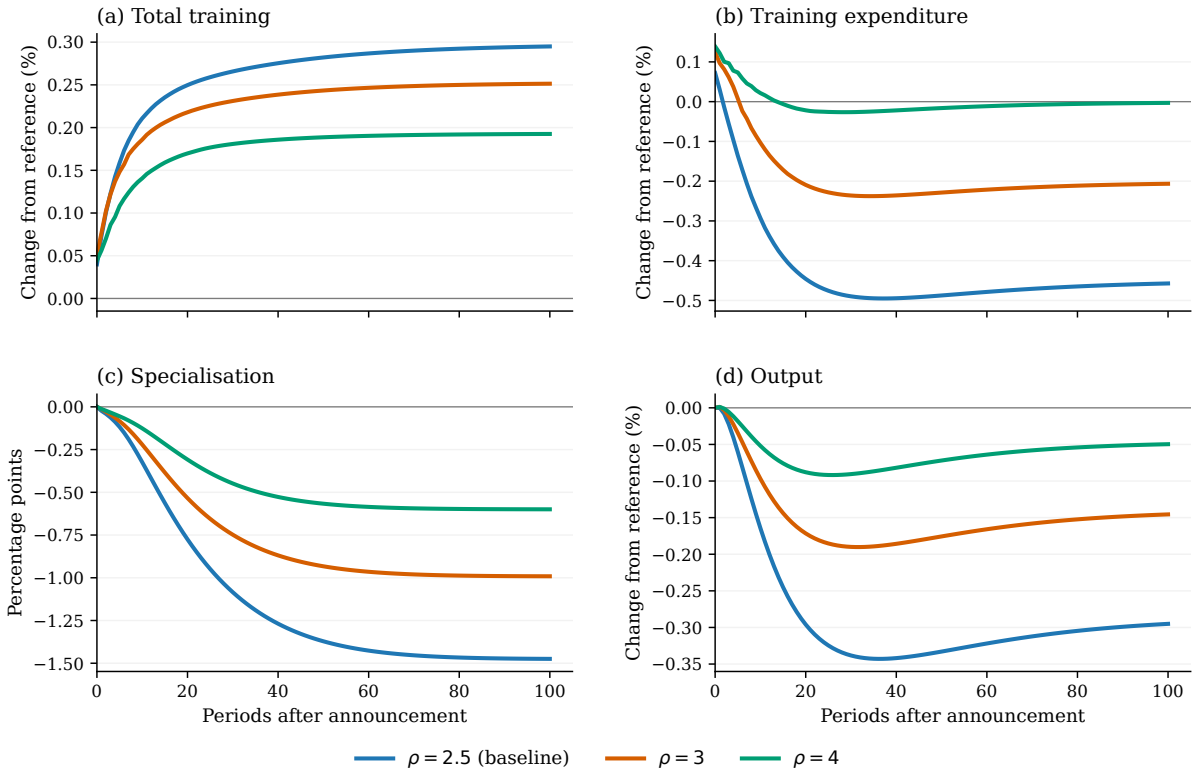



Figure 11: Training and skill portfolios under greater forecast uncertainty

*Notes:* Each curve compares the policy chosen under a permanent 20% increase in conditional level variances with the original policy, applying both to unchanged reference return histories. Expected level-return paths are the same. All three curvatures start from the same baseline population, conditional on middle-state common returns. Marginal training cost is matched at the reference investment, with $\sigma = 2$ throughout. Panels (a), (b) and (d) show percentage changes in mean training quantity $I_1 + I_2$, expenditure $K$ and output $Y$, relative to each curvature's evolving reference. Panel (c) shows the change in mean individual specialisation in percentage points. Means integrate over common and individual histories.

**The role of training costs.** Higher training-cost curvature reduces the medium-run response to uncertainty. After twenty periods, output is 0.30%, 0.17% and 0.09% below its respective reference as $\rho$ rises from 2.5 to 3 and 4. Specialisation falls by 0.78, 0.53 and 0.31 percentage points, while total training quantity rises by 0.25%, 0.22% and 0.17%. Hence, higher curvature reduces the increase in total training and therefore the subsequent change in skill composition. The smaller output loss is due to less portfolio adjustment, rather than a larger increase in total skills. Since curvature changes desired portfolios and cost levels as well as the transition, this exercise does not demonstrate the effects of differences in skill adjustment speed.

**Risk aversion and skill portfolios.** The skill diversification response to increased uncertainty also appears under risk neutrality. Figure 12 repeats the uncertainty comparison for $\gamma = 0, 1, 2$. Within each row, policies are chosen under different forecast variances and evaluated under the same realised return histories. Raising conditional variances by 20% reduces long-run mean specialisation by 0.84, 1.47 and 1.74 percentage points, respec-

tively. Hence, risk aversion is not necessary to generate a skill diversification response to increased uncertainty. Irrespective of risk aversion, workers under increased uncertainty benefit more from being able to use different skills as returns change ("wait and see" approach).

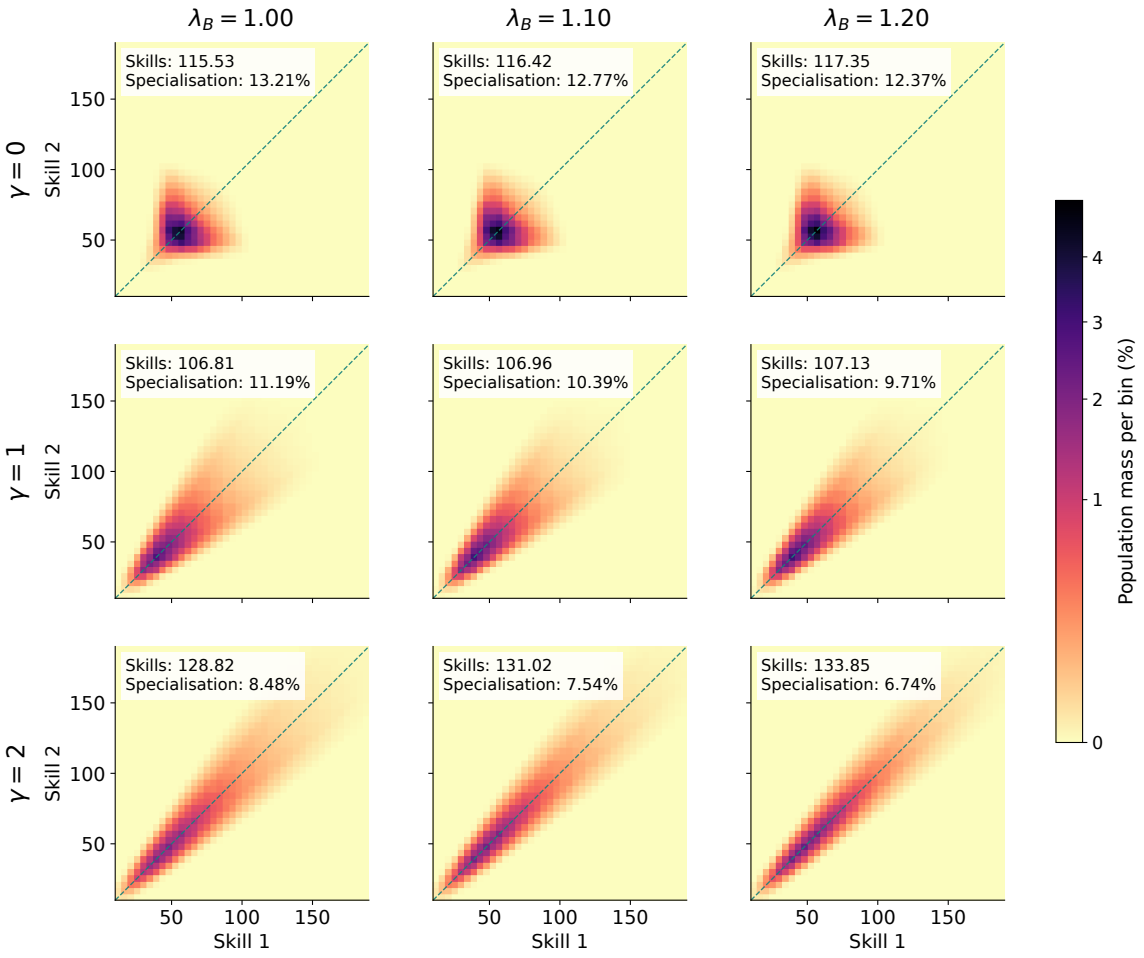



Figure 12: Uncertainty and the distribution of worker skills

*Notes:* Rows vary CRRA curvature; columns multiply each common return's conditional level variance by $\lambda_B = 1, 1.10, 1.20$. Expected return paths are unchanged. Policies are chosen under their respective forecasts and evaluated under the same reference histories. Each row shares a common initial population across columns, but populations differ across rows. Panels show long-run distributions averaged over common histories. Colours indicate population mass with common scales, and the dashed line marks perfectly balanced skills.

Total skills rise by 1.57%, 0.30% and 3.90% at $\gamma = 0, 1, 2$, respectively. Unlike the log-utility case, long-run output rises at $\gamma = 0$ and 2, by 1.30% and 2.93%. The output decline under log utility is therefore not a general result. Changing $\gamma$ also changes intertemporal substitution, so these comparisons do not separately identify the motives for diversification.

The uncertainty experiment suggests that less predictable returns can encourage workers to build a more balanced skill portfolio. This prediction concerns the content of training

– the direction of training in the skill space – which our empirical indicators do not capture. Training for future employment, as recorded in our data, could either deepen existing specialisation or broaden workers' skills. Distinguishing these responses requires information on the skills acquired through training or subsequently used at work. The training records in our data do not provide enough information on the purpose of training for us to infer the acquired skills from it. We plan to instead estimate the acquired skills from workers switching occupations in future work, under the assumption that they gained those skills in the preceding training.

## 6.4 Training and the Gains from Technological Change

**Additional training after a permanent skill return increase.** Training allows workers to benefit more from technological change. In the baseline, allowing investment to respond increases the discounted cumulative output gain from a permanent return improvement by 25.8% through period twenty. Raising training-cost curvature from $\rho = 2.5$ to $\rho = 4$ reduces this amplification to 13.6%, as workers are less able to quickly adapt their skill portfolios to utilise the increased skill return.

Using the original log-return process, we raise the return to skill 1, $B_1$, by 10% at impact and keep it 10% higher along every subsequent history. Workers know the improvement is permanent. Transition probabilities, $B_2$ and individual productivity are unchanged. Both cost specifications start from the same population after a common-history burn-in, conditional on middle-state common skill-neutral returns. We then subtract their own path without the $B_1$ increase.

The costs use the investment cost normalisation in Section 6.3, with $\kappa_{2.5} = 1$ and $\kappa_4 \simeq 0.13$, matching marginal cost at $\bar{I} \simeq 2.81$. Other parameters are unchanged.

Figure 13 compares three paths. The blue line, endogenous training, allows training to respond to higher returns. The orange line, unchanged training policy, holds fixed the investment choices made without the $B_1$ improvement along the same underlying return histories – i.e., workers continue to train, and these training policies remain affordable because output rises at given stocks due to increased $B_1$. The grey line holds initial skill stocks fixed. Workers optimise their skill-use weights along every path.

The responses coincide at impact because stocks are predetermined. They diverge as training shifts towards skill 1. After twenty periods, output is 7.55% above its reference at $\rho = 2.5$ and 6.43% above it at $\rho = 4$, compared with about 5.22% in both cases when training choices are unchanged. After one hundred periods, endogenous-training gains reach 8.92% and 6.94%, while unchanged-training gains remain around 5.1%.

Discounting through period twenty at $\beta = 0.99$, cumulative output gains are 98.76 and

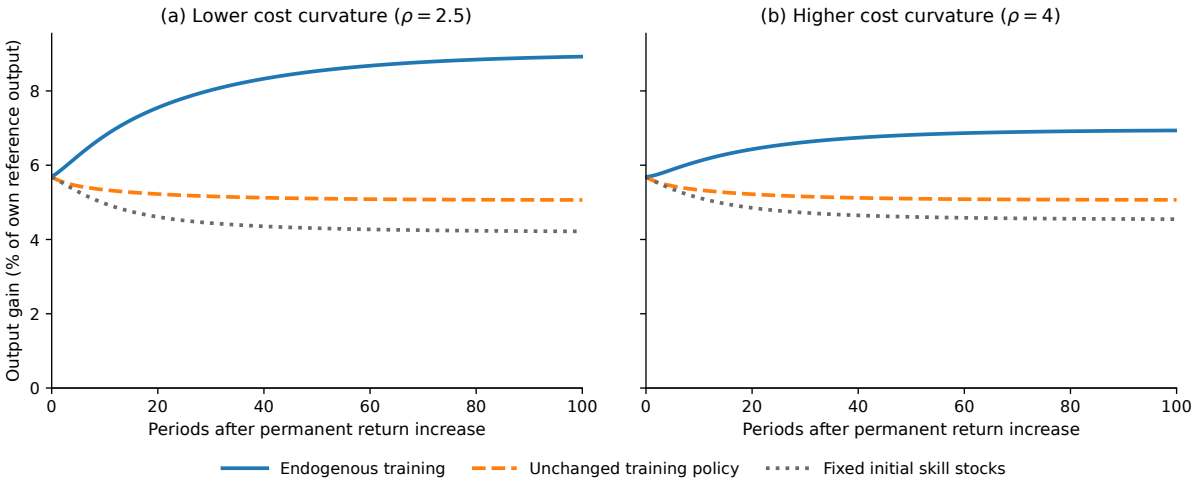



Figure 13: The output gains from training after a permanent return improvement

*Notes:* The common return to skill 1 is permanently 10% higher from period zero. Panels (a) and (b) use $\rho = 2.5$ and $\rho = 4$, with marginal training cost matched at the reference investment. Blue allows training to respond; orange retains the investment path without the return increase; grey fixes initial skill stocks. Skill-use weights adjust in every case. Each curve subtracts its matched path without the increase. All three curves within a panel divide by that cost specification's output without the increase. Initial workers and underlying return histories are common across panels. Means integrate over common and individual histories.

86.64 units per worker at the lower and higher curvatures. Without a training response, they are 78.50 and 76.28 units. Additional training accounts for 20.5% and 12.0% of the respective gains. Comparing with fixed stocks instead gives gains 34.4% and 17.9% larger, but also includes the interaction of ordinary skill accumulation with higher returns.

More convex costs (greater value of $\rho$) reduce the output contribution of training, but also reduce resources spent on it because workers respond by increasing training less. Through period twenty, the expenditure gain falls from 41.24 to 21.24 units per worker, while the consumption gain rises from 57.52 to 65.40 units.

**Permanent skill return increases under greater uncertainty.** We finally compare the a permanent skill-1 return increase with the same increase accompanied by also greater forecast uncertainty. Both paths use the same reference without either change. Their gap therefore includes the effect uncertainty would have even without a return improvement.

We combine the permanent 10% increase in $B_1$ with a permanent 20% increase in conditional variances that preserves expected return paths. This requires the matched process from Section 6.3, so we recompute the permanent-return response under that process. We retain the initial workers from Figure 13 and evaluate both policies on identical realised histories, including the permanent return increase.

Increased uncertainty lowers the cumulative output response to the skill-1 return improvement in Figure 14. Through period twenty, the discounted gain falls from 92.72 to 90.11

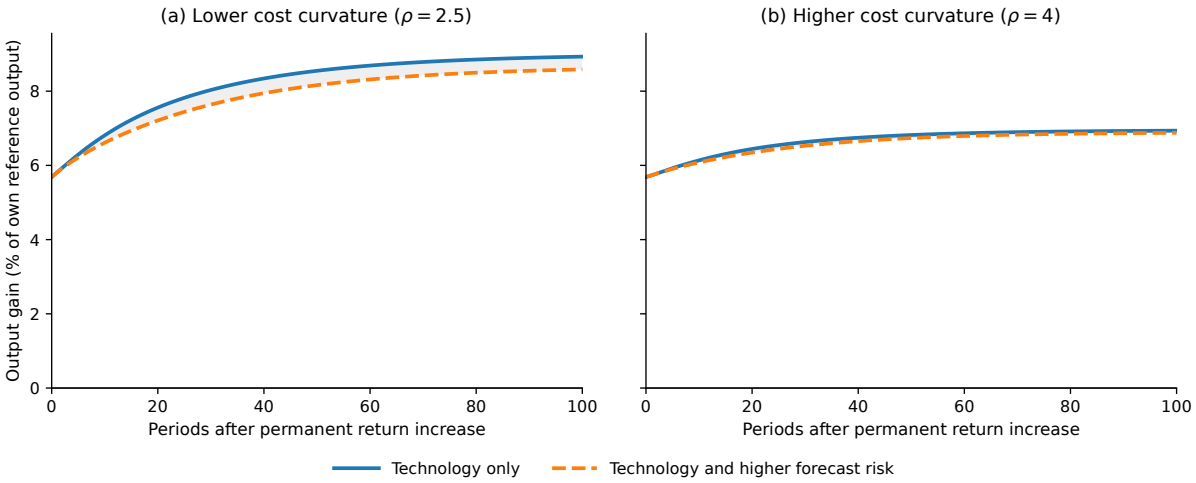



Figure 14: Permanent return improvements accompanied by greater forecast uncertainty

*Notes:* Blue lines show output gains from a permanent 10% increase in $B_1$. Orange lines use policies chosen under an additional permanent 20% increase in each common return's conditional level variance, with expected return paths unchanged. All policies are evaluated under the same realised reference histories. Both curves subtract the reference without either change and divide by its output at each date. Their gap therefore includes the effect of uncertainty that would also occur without the return improvement. The two panels retain the investment cost normalisations and common initial workers from Figure 13.

units per worker at $\rho = 2.5$, and from 82.12 to 81.41 at $\rho = 4$. These are reductions of 2.8% and 0.9%, respectively. Through period one hundred, the reductions are 4.1% and 1.1%.

Intuitively, higher investment cost curvature means that workers adjust their skill investment less, both in response to both higher returns and in response to increased uncertainty. The gains from skill adjustment are therefore smaller, both with and without the additional increase in uncertainty. The difference between the two uncertainty scenarios is smaller too. Even when uncertainty is greater, the cumulative output gain at $\rho = 2.5$ exceeds the output gain at $\rho = 4$ under normal levels of uncertainty. These results suggest that uncertainty about future skill returns dampens the additional output gain from higher realized skill returns, as workers make less use of the increased skill returns.

# 7 Conclusion

This paper studies how workers adjust their human capital in response to skill-level technological change. We construct a new measure of technological exposure by linking patent texts to O*NET skill descriptors and aggregating it using occupational skill requirements. Combining this measure with UK longitudinal worker data allows us to trace how wages, training and employer mobility evolve with exposure. We find that technological change raises hourly wages on average, consistent with net complementarity between technology and labour, and that workers respond through forward-looking investment in skills and

voluntary moves to better jobs.

The wage, hours and unemployment responses suggest that technological change primarily raises the marginal product of labour. Earnings increase mainly through hourly pay, while there is little adjustment in hours and no clear increase in unemployment. The larger wage gains among workers in non-routine occupations and those with higher education are consistent with established findings on technological change and wage inequality.

The training and mobility responses show how workers adjust to changing technological opportunities. Workers train more, particularly to prepare for future jobs, and make more voluntary changes of employer for promotion or a better job. Taken together, these findings suggest that workers develop their skills in anticipation of future returns and seek jobs that better reward those skills. Our results thus highlight the forward-looking nature of human-capital adjustment, both through investment in skills and through changes in where those skills are used.

Our model provides an explanation for the joint increase in wages and training. Higher skill returns raise output and provide more resources for both consumption and skill investment. In our numerical exercises, these additional current resources account for most of the increase in training expenditure, while changes in future payoffs mainly direct investment towards the skill whose return improves. The resulting skill accumulation allows workers to gain more from technological change. In an illustrative exercise, allowing training to respond to a permanent improvement in one skill's return increases the discounted cumulative output gain by about a quarter at a horizon of twenty model periods, relative to retaining the investment path workers would have followed without the return improvement. More convex training costs reduce this amplification.

Uncertainty also affects which skills workers choose to acquire. Holding expected return paths and realised return histories fixed, greater forecast uncertainty leads workers to build more balanced skill portfolios in our numerical exercises, with only a small change in total training in the baseline. This response also occurs under risk neutrality. Holding both skills allows workers to make greater use of whichever skill turns out to pay more. Our results therefore suggest that uncertainty about future skill demand can affect the skills workers choose to acquire and therefore the aggregate skill distribution.

A next step is to connect the model more closely to occupational skill requirements and workers' training decisions. Extending it to three or four skills would allow a more direct mapping to categories such as mathematical, verbal and social skills. Together with a mapping from technological exposure to skill returns, this would allow us to quantify how the variance and co-movement of returns affect training and specialisation. Furthermore, by introducing frictions in workers' occupational choice we would be able to study how

these investment decisions depend on the jobs available to workers. Compared to the frictionless occupation choice in our current model, workers may in practice be unwilling to acquire skills that are valuable (including from a macroeconomic perspective) but that are difficult to use in the occupations available to them. Such frictions could also reduce the value of holding a balanced vs. a specialised skill portfolio, since workers then cannot freely change how they use their skills. We would use this setting to assess how training subsidies and insurance against losses in skill returns affect workers' skill choices, output and welfare.

# References


Acemoglu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics* (pp. 1043–1171). Elsevier.

Acemoglu, D., Kong, F., & Restrepo, P. (2024). Tasks At Work: Comparative Advantage, Technology and Labor Demand.

Acemoglu, D., & Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, *128*(6), 2188–2244.

Akerman, A., Gaarder, I., & Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics*, *130*(4), 1781–1824.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration*. *The Quarterly Journal of Economics*, *118*(4), 1279–1333.

Bartel, A., Ichniowski, C., & Shaw, K. (2007). How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills. *The Quarterly Journal of Economics*, *122*(4), 1721–1758.

Bartel, A. P., & Sicherman, N. (1998). Technological Change and the Skill Acquisition of Young Workers. *Journal of Labor Economics*, *16*(4), 718–755.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of political economy*, *75*(4, Part 1), 352–365.

Blundell, R., Costa Dias, M., Goll, D., & Meghir, C. (2021). Wages, Experience, and Training of Women over the Life Cycle. *Journal of Labor Economics*, *39*(S1), S275–S315.

Cavounidis, C., & Lang, K. (2020). Ben-porath meets lazear: Microfoundations for dynamic skill formation [Publisher: University of Chicago Press]. *Journal of Political Economy*, *128*(4), 1405–1435.

Gathmann, C., Grimm, F., & Winkler, E. (2024). *AI, Task Changes in Jobs, and Worker Reallocation* (IZA Discussion Paper No. 17554). IZA Institute of Labor Economics. https://www.iza.org/publications/dp/17554/ai-task-changes-in-jobs-and-worker-reallocation

Gathmann, C., & Schoenberg, U. (2010). How General Is Human Capital? A Task-Based Approach. *Journal of Labor Economics*, *28*(1).

Goos, M., & Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *The Review of Economics and Statistics*, *89*(1), 118–133.

Graetz, G., & Michaels, G. (2018). Robots at Work. *The Review of Economics and Statistics*, *100*(5), 753–768.

Heß, P., Janssen, S., & Leber, U. (2023). The effect of automation technology on workers' training participation. *Economics of Education Review*, *96*, Article 102438.

Jolivet, G., & Postel-Vinay, F. (2025). A Structural Analysis of Mental Health and Labour Market Trajectories. *The Review of Economic Studies*, *92*(3), 1920–1954.

Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, *95*(1), 161–182.

Katz, L. F., & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics*, *107*(1), 35–78.

Kelly, B., Papanikolaou, D., Seru, A., & Taddy, M. (2021). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*, *3*(3), 303–20.

Kogan, L., Papanikolaou, D., Schmidt, L., & Seegmiller, B. (2019). *Technology, Vintage-Specific Human Capital, and Labor Displacement: Evidence from Linking Patents with Occupations*.

Kogan, L., Papanikolaou, D., Schmidt, L. D. W., & Seegmiller, B. (2023). *Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data* (Working Paper No. 31846) [Revised November 2025]. National Bureau of Economic Research.

Lazear, E. P. (2009). Firm-Specific Human Capital: A Skill-Weights Approach. *Journal of Political Economy*, *117*(5).

Lise, J., & Postel-Vinay, F. (2020). Multidimensional Skills, Sorting, and Human Capital Accumulation. *American Economic Review*, *110*(8), 2328–2376.

Merton, R. C. (1969). Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics*, *51*(3), 247–257. Retrieved May 15, 2025, from http://www.jstor.org/stable/1926560

Michaels, G., Natraj, A., & Van Reenen, J. (2014). Has ICT Polarized Skill Demand? Evidence from Eleven Countries over Twenty-Five Years. *The Review of Economics and Statistics*, *96*(1), 60–77.

Postel-Vinay, F., & Sepahsalari, A. (2023). Labour Mobility and Earnings in the UK, 1992–2017. *The Economic Journal*, *133*(656), 3071–3098.

Prytkova, E., Fabien, P., Deyu, L., Sugat, C., & Tommaso, C. (2024, October). *The Employment Impact of Emerging Digital Technologies*.

Silos, P., & Smith, E. (2015). Human capital portfolios. *Review of Economic Dynamics*, *18*(3), 635–652.

University of Essex, Institute for Social and Economic Research. (2026). Understanding Society: Waves 1–15, 2009–2024 and Harmonised BHPS: Waves 1–18, 1991–2009 [21st edition. Data collection. Study Number 6614].

Webb, M. (2019, November 6). The Impact of Artificial Intelligence on the Labor Market.

# A Appendix Figures

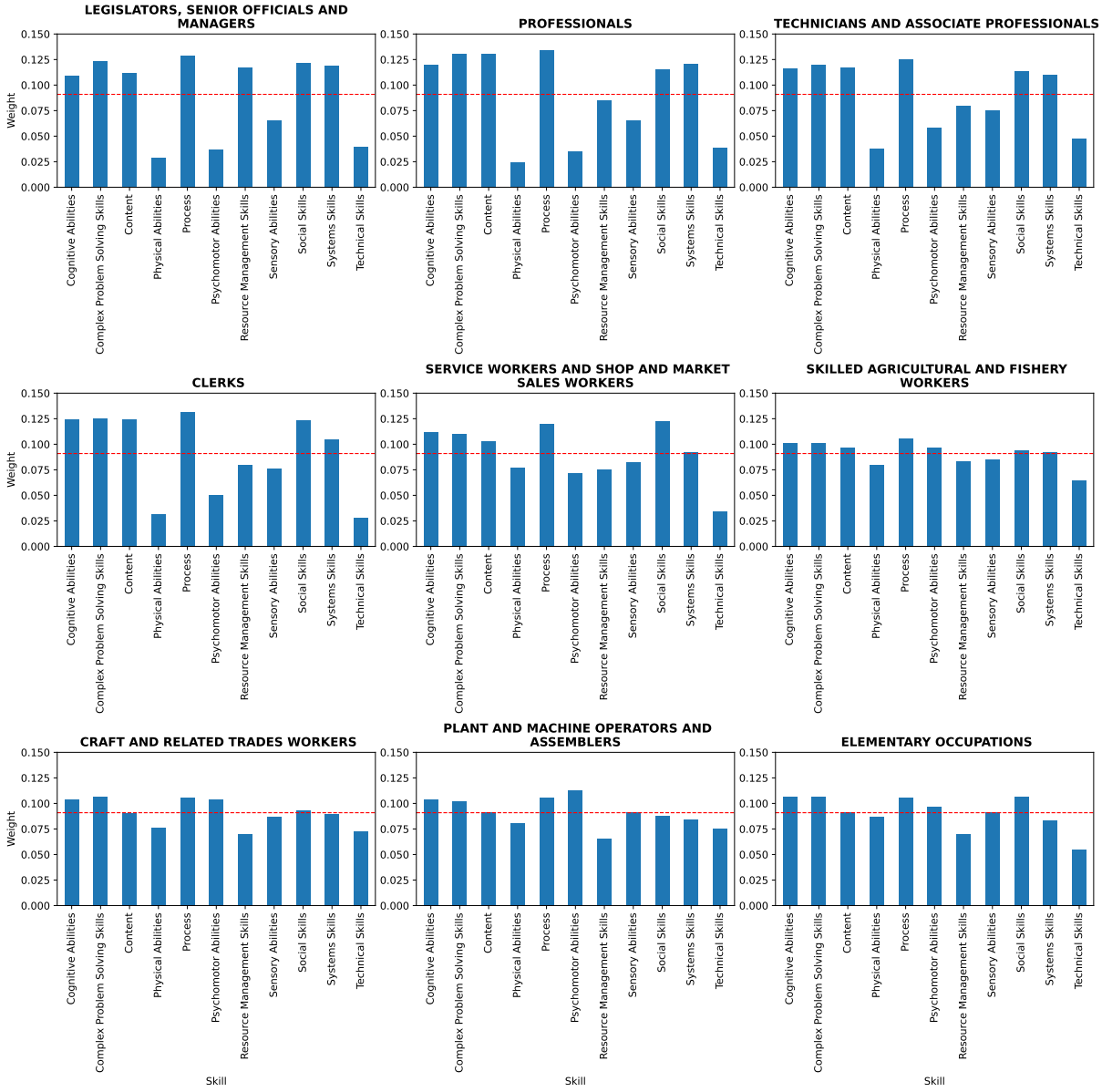



Figure 15: Relative skill weights across ISCO-88 major occupation groups

*Notes:* Red dashed line indicates equal weight across 11 O*NET broad skill categories

# B Empirical Appendix

This appendix documents the occupational classification used in the wage comparisons and presents supplementary results for hourly pay, training purposes, and job mobility. The estimates use the specification in equation (3), with exposure assigned using occupation at baseline. The figures report responses to a one-standard-deviation increase in residualised occupational exposure and pointwise 95% confidence intervals.

## B.1 Occupational Classification

Table B1 reports the routine-task classification underlying panel (a) of figure 5. Clerks, craft and related trades workers, and plant and machine operators and assemblers are classified as routine. The remaining major groups are classified as non-routine. Classification uses the worker's occupation at the exposure baseline; subsequent occupation changes do not reassign the group for that response window. This is a broad grouping of occupations rather than an individual measure of tasks performed.

Table B1: Occupational classification by routine-task group

| ISCO-88 major group | Occupational group | Classification |
|---|---|---|
| 1 | Legislators, senior officials and managers | Non-routine |
| 2 | Professionals | Non-routine |
| 3 | Technicians and associate professionals | Non-routine |
| 4 | Clerks | Routine |
| 5 | Service workers and shop and market sales workers | Non-routine |
| 6 | Skilled agricultural and fishery workers | Non-routine |
| 7 | Craft and related trades workers | Routine |
| 8 | Plant and machine operators and assemblers | Routine |
| 9 | Elementary occupations | Non-routine |

*Notes:* The table reproduces the grouping used in the wage regressions. Major groups are derived from the baseline three-digit ISCO-88 code; routine occupations comprise groups 4, 7, and 8. Armed forces and missing or invalid codes are excluded from this classification. Group membership is assigned separately at each exposure baseline and retained over the corresponding response window. Occupational titles follow the International Labour Organization's ISCO-88 structure.

## B.2 Hourly-Pay Responses by Sex

Figure B1 supplements the occupational and educational comparisons in figure 5 with separate hourly-pay profiles for men and women. The male profile becomes more positive at longer horizons, while the female estimates are smaller and less precise. At horizon ten, the estimates are approximately 5.3% and 0.4%, respectively. These profiles describe the observed subgroups. Selection into occupations and into the observed wage sample prevents a causal interpretation of differences between them.

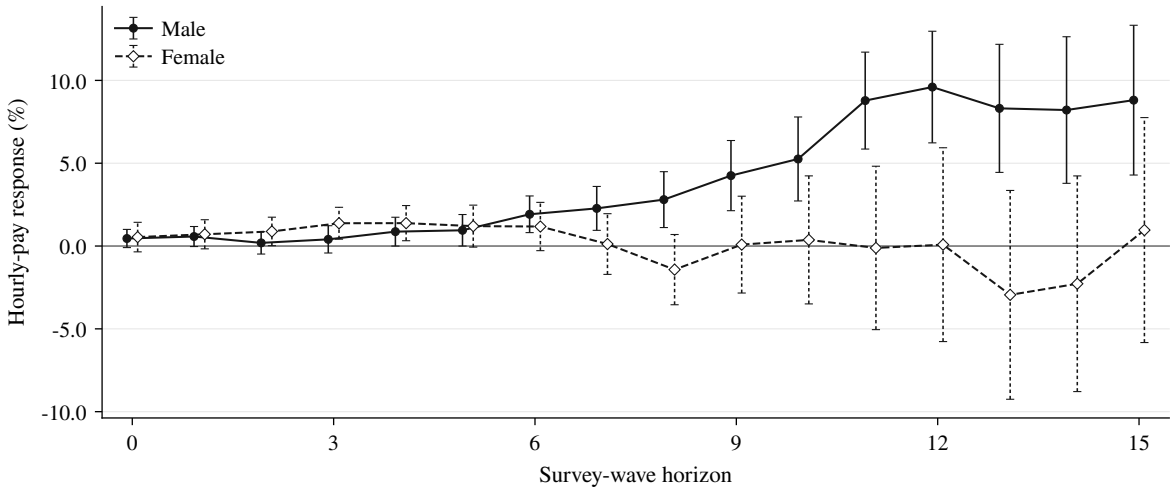



Figure B1: Hourly-pay responses by sex

*Notes:* The outcome is the change in log real hourly pay from the wave before exposure to the indicated horizon, multiplied by 100. Sex is recorded at baseline. Each profile is estimated separately and uses the same exposure scale as the main results. Regressions include individual and baseline-occupation fixed effects and calendar-year fixed effects for the baseline interview, use longitudinal weights at the outcome horizon, and cluster standard errors by baseline occupation–exposure-year. Bars show pointwise 95% confidence intervals. Samples vary across groups and horizons. The profiles are descriptive subgroup comparisons; the figure does not report a test of equality between them.

## B.3 Additional Training Purposes

Figure B2 supplements the current-job and future-job training outcomes in figure 7 with four additional reported purposes. Promotion-related training has a positive response at later horizons: at horizon eight, the estimate is 2.0 additional positive reports per 100 workers, with a 95% confidence interval of 0.6 to 3.5. Training for promotion provides a further forward-looking margin of skill investment, consistent with the preparation for future opportunities discussed in the main text and complementary to the voluntary employer moves in figure 8.

Starting-job training and leisure training show little sustained response. Training to maintain professional status is reported separately and does not display a sustained positive profile. These comparisons describe patterns across the reported purposes without combining them into a single measure of forward-looking investment.

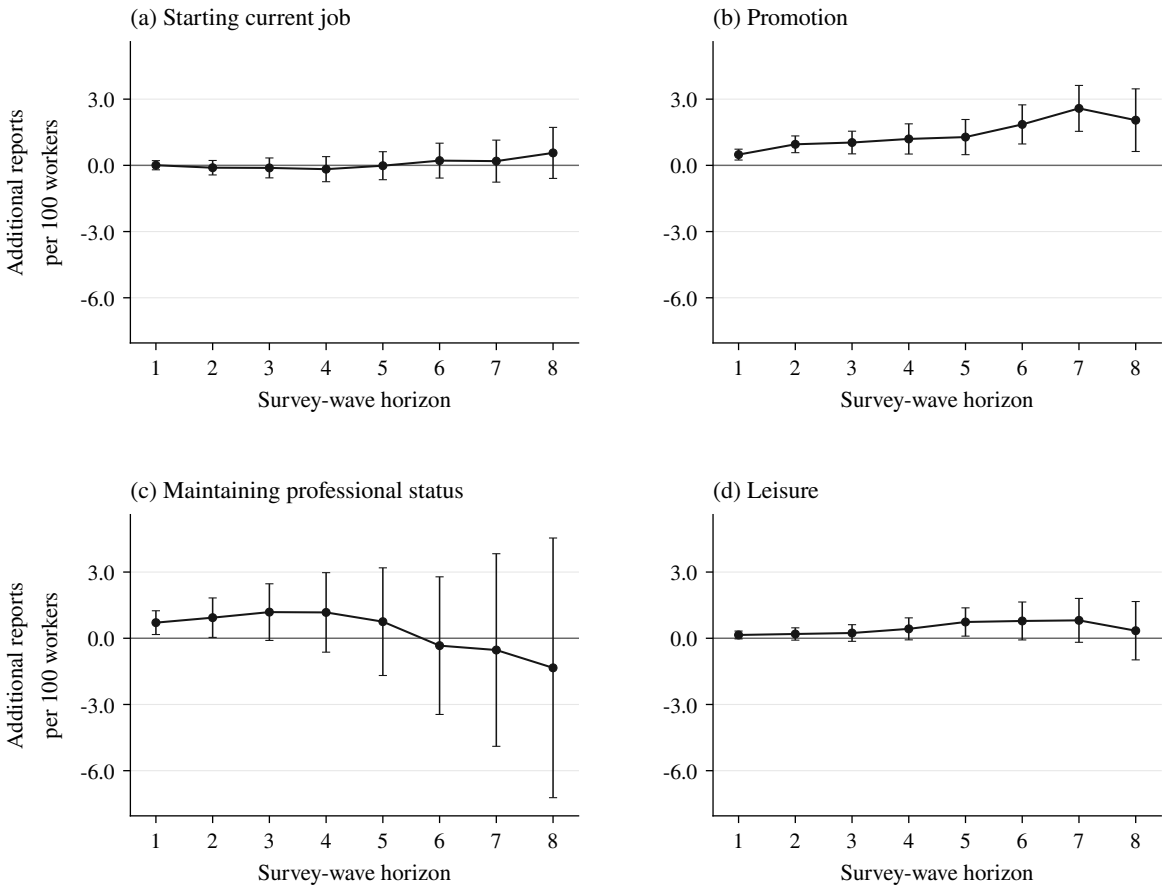



Figure B2: Additional purposes of training

*Notes:* Each outcome counts future waves with a positive report of the indicated training purpose over the first $h$ future waves after exposure, with at most one contribution per wave. Responses are additional positive reports per 100 workers; the panels use a common vertical scale. Purpose outcomes use Understanding Society only, with baseline waves beginning in Wave 2 and future outcomes extending through the wave assigned to 2022. Every wave in the future window must have a known outcome. Known non-trainees contribute zero; unknown purposes are not coded as zero. A course can have several purposes, and samples differ by purpose and horizon, so the panels do not form an additive decomposition. Counts measure positive reporting waves, rather than distinct courses. Estimation follows figure B1.

## B.4 Additional Measures of Job Mobility

The main text focuses on employer changes and reported reasons for moving. Figure B3 places that outcome alongside broader measures of job and occupation mobility. A job change is an identified employer change or an occupation change within a continuing employer. Within-employer mobility therefore refers specifically to changes of occupation; it does not capture every change of position or responsibility. Occupation mobility records changes between the main occupations observed at successive interviews, including changes that accompany an employer move.

The point estimates in all four panels are positive at longer horizons, although uncertainty is substantial for within-employer occupation changes. These results provide descriptive context for the employer-mobility analysis. Their relative magnitudes cannot be read as contributions to a total: the outcomes overlap and have different observation requirements. In particular, the occupation-related panels have shorter survey coverage than employer changes.

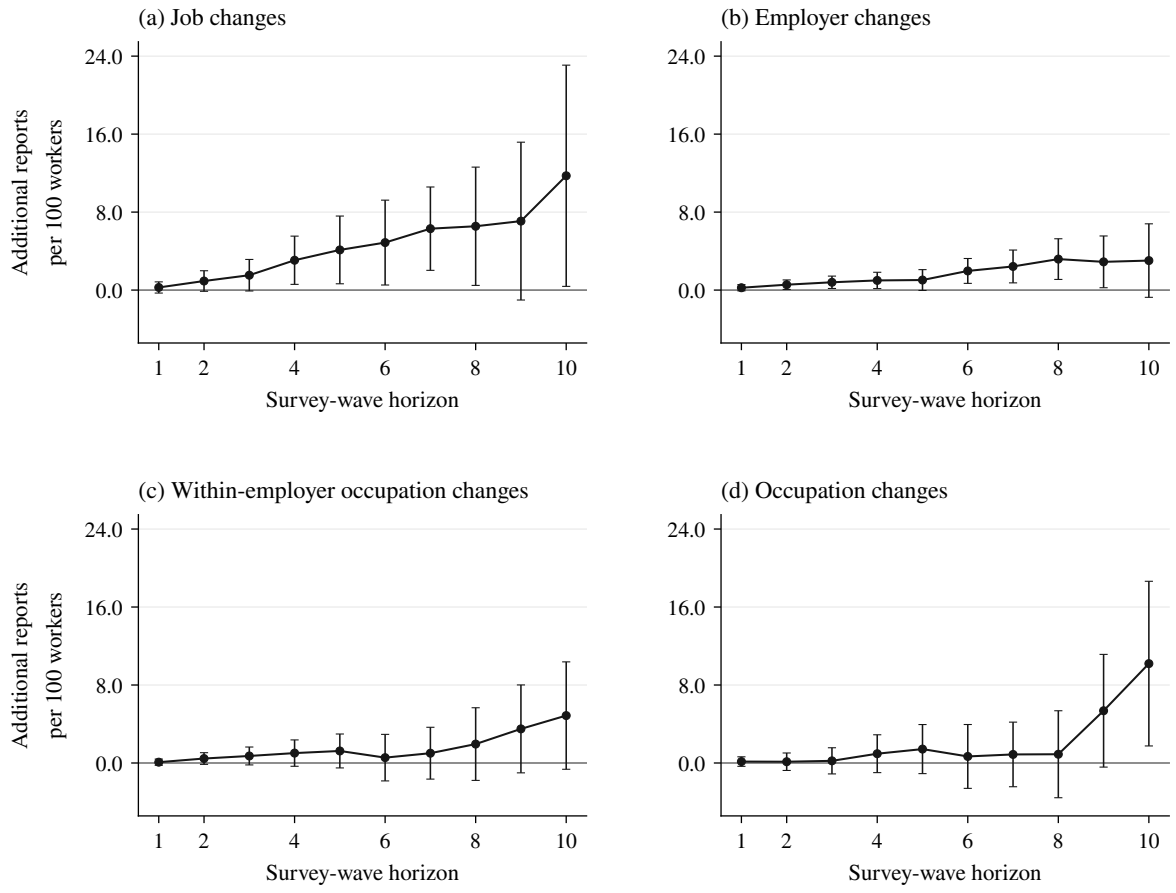



Figure B3: Job, employer, and occupation mobility

*Notes:* Each outcome counts future waves containing the indicated change over the first $h$ future waves after exposure, with at most one contribution per wave and a fully observed future outcome window. Responses are additional positive reports per 100 workers; the panels use a common vertical scale. Employer changes use annual employment histories and include moves with intervening nonemployment. Occupation-related outcomes require observed occupations at successive interviews; within-employer changes additionally require evidence of a continuing employer. At horizon ten, sample sizes are 4,038, 25,389, 5,280, and 5,938 individual-wave observations in panels (a)–(d), respectively. The outcomes overlap and are not an additive decomposition; counts are reporting waves rather than the number of underlying transitions. Estimation follows figure B1.

## B.5 Estimates at Selected Horizons

Tables B2–B5 report the coefficients, clustered standard errors, and estimation sample sizes underlying the figures at horizons five, eight, ten, and fifteen, where available.

Table B2: Selected estimates for pay, hours, and unemployment

| | Horizon (survey waves) | | | |
|---|---|---|---|---|
| | 5 | 8 | 10 | 15 |
| Annualised usual earnings | 0.58 | 0.58 | 1.64 | 6.97 |
| | (0.47) | (0.84) | (1.39) | (2.40) |
| *N* | 94,190 | 57,370 | 40,321 | 17,282 |
| Hourly pay | 1.00 | 1.03 | 3.07 | 4.86 |
| | (0.38) | (0.64) | (1.08) | (1.77) |
| *N* | 90,653 | 55,168 | 38,680 | 16,506 |
| Weekly hours | −0.49 | −0.90 | −1.39 | 0.82 |
| | (0.38) | (0.66) | (1.01) | (1.82) |
| *N* | 90,827 | 55,278 | 38,764 | 16,532 |
| Unemployment reports | 0.08 | 0.00 | −0.79 | 0.36 |
| | (0.25) | (0.49) | (0.88) | (2.08) |
| *N* | 141,936 | 84,679 | 53,713 | 7,558 |

*Notes:* Pay and hours are changes in logs multiplied by 100 (approximate percent); unemployment is additional positive reports per 100 workers. Responses are to a one-standard-deviation increase in occupational exposure. Standard errors in parentheses are clustered by baseline occupation–exposure-year. All specifications include individual and baseline-occupation fixed effects and calendar-year fixed effects for the baseline interview, and use longitudinal survey weights at the outcome horizon. *N* is the number of individual–baseline-wave observations; samples vary across outcomes, subgroups, and horizons. A dash denotes a horizon outside the reported profile.

Table B3: Selected hourly-pay estimates by subgroup

| | Horizon (survey waves) | | | |
|---|---|---|---|---|
| | 5 | 8 | 10 | 15 |
| Non-routine occupations | 2.44 | 2.65 | 5.77 | 7.27 |
| | (0.62) | (1.18) | (2.05) | (3.44) |
| *N* | 62,908 | 37,385 | 25,642 | 10,488 |
| Routine occupations | −1.20 | −2.00 | −2.55 | −2.52 |
| | (0.62) | (0.98) | (1.46) | (2.60) |
| *N* | 26,150 | 16,563 | 12,083 | 5,535 |
| Higher education | 2.06 | 2.46 | 5.99 | 11.82 |
| | (0.82) | (1.24) | (2.21) | (4.38) |
| *N* | 36,308 | 21,146 | 14,055 | 5,328 |
| No higher education | 0.03 | −0.49 | 1.62 | 1.05 |
| | (0.46) | (0.75) | (1.21) | (2.01) |
| *N* | 53,621 | 33,611 | 24,285 | 11,010 |
| Men | 0.95 | 2.80 | 5.26 | 8.81 |
| | (0.49) | (0.86) | (1.29) | (2.31) |
| *N* | 40,905 | 25,008 | 17,702 | 7,801 |
| Women | 1.20 | −1.42 | 0.37 | 0.96 |
| | (0.65) | (1.08) | (1.97) | (3.46) |
| *N* | 49,687 | 30,123 | 20,947 | 8,691 |

*Notes:* Hourly-pay changes in logs multiplied by 100 (approximate percent). Groups are defined at baseline. These are descriptive subgroup comparisons. Responses are to a one-standard-deviation increase in occupational exposure. Standard errors in parentheses are clustered by baseline occupation–exposure-year. All specifications include individual and baseline-occupation fixed effects and calendar-year fixed effects for the baseline interview, and use longitudinal survey weights at the outcome horizon. *N* is the number of individual–baseline-wave observations; samples vary across outcomes, subgroups, and horizons. A dash denotes a horizon outside the reported profile.

Table B4: Selected estimates for training and its purposes

| | Horizon (survey waves) | | | |
|---|---|---|---|---|
| | 5 | 8 | 10 | 15 |
| Any training | 5.38 | 11.56 | 18.51 | — |
| | (0.82) | (1.66) | (2.95) | — |
| *N* | 142,873 | 84,351 | 53,041 | — |
| Current-job skills | 0.15 | 2.14 | — | — |
| | (1.41) | (3.14) | — | — |
| *N* | 73,012 | 37,144 | — | — |
| Future-job opportunities | 1.21 | 4.15 | — | — |
| | (0.72) | (1.27) | — | — |
| *N* | 66,166 | 32,495 | — | — |
| Starting current job | −0.02 | 0.56 | — | — |
| | (0.32) | (0.59) | — | — |
| *N* | 65,210 | 31,896 | — | — |
| Promotion | 1.28 | 2.05 | — | — |
| | (0.41) | (0.72) | — | — |
| *N* | 65,106 | 31,852 | — | — |
| Maintaining professional status | 0.75 | −1.34 | — | — |
| | (1.24) | (3.00) | — | — |
| *N* | 72,206 | 36,447 | — | — |
| Leisure | 0.74 | 0.34 | — | — |
| | (0.33) | (0.67) | — | — |
| *N* | 64,856 | 31,738 | — | — |

*Notes:* Coefficients and standard errors are additional positive reports per 100 workers, with at most one contribution per worker in each future wave. Purposes can overlap. Any training is reported through horizon ten and purpose-specific outcomes through horizon eight. Responses are to a one-standard-deviation increase in occupational exposure. Standard errors in parentheses are clustered by baseline occupation–exposure-year. All specifications include individual and baseline-occupation fixed effects and calendar-year fixed effects for the baseline interview, and use longitudinal survey weights at the outcome horizon. *N* is the number of individual–baseline-wave observations; samples vary across outcomes, subgroups, and horizons. A dash denotes a horizon outside the reported profile.

Table B5: Selected estimates for job and employer changes

| | Horizon (survey waves) | | | |
|---|---|---|---|---|
| | 5 | 8 | 10 | 15 |
| Job changes | 4.12 | 6.55 | 11.73 | — |
| | (1.77) | (3.10) | (5.79) | — |
| *N* | 40,063 | 14,174 | 4,038 | — |
| Employer changes | 1.03 | 3.18 | 3.03 | — |
| | (0.54) | (1.06) | (1.92) | — |
| *N* | 90,816 | 45,554 | 25,389 | — |
| Within-employer occupation changes | 1.23 | 1.94 | 4.86 | — |
| | (0.89) | (1.90) | (2.81) | — |
| *N* | 48,544 | 17,820 | 5,280 | — |
| Occupation changes | 1.43 | 0.90 | 10.20 | — |
| | (1.28) | (2.28) | (4.31) | — |
| *N* | 51,978 | 19,542 | 5,938 | — |
| Voluntary employer changes | 1.36 | 3.72 | 4.38 | — |
| | (0.42) | (0.81) | (1.42) | — |
| *N* | 89,557 | 44,718 | 24,797 | — |
| Involuntary employer changes | −0.11 | 1.02 | 0.71 | — |
| | (0.23) | (0.50) | (0.95) | — |
| *N* | 89,086 | 44,438 | 24,689 | — |

*Notes:* Coefficients and standard errors are additional positive reports per 100 workers. Outcomes overlap and have different samples; they do not form an additive decomposition. Profiles are reported through horizon ten. Responses are to a one-standard-deviation increase in occupational exposure. Standard errors in parentheses are clustered by baseline occupation–exposure-year. All specifications include individual and baseline-occupation fixed effects and calendar-year fixed effects for the baseline interview, and use longitudinal survey weights at the outcome horizon. $N$ is the number of individual–baseline-wave observations; samples vary across outcomes, subgroups, and horizons. A dash denotes a horizon outside the reported profile.

# C Model Appendix

## C.1 Skill Use and Investment

Given positive effective skills $A_iS_i$, the worker chooses $J$ to maximise $\sum_i A_iJ_iS_i$ subject to $\sum_i J_i^\sigma \leq 1$. The constraint binds. With multiplier $\lambda$, the first-order conditions are $A_iS_i = \lambda\sigma J_i^{\sigma-1}$, giving

$$J_i^*(S,A) = \frac{(A_iS_i)^{1/(\sigma-1)}}{\left[(A_1S_1)^q + (A_2S_2)^q\right]^{1/\sigma}}, \qquad q = \frac{\sigma}{\sigma-1}. \tag{14}$$

The objective is linear and the feasible set is convex, so these conditions give the global maximum. Substituting the weights into production gives (7), with marginal product and return elasticity

$$Y_{S_i} = Y^{1-q}A_i^qS_i^{q-1}, \qquad \frac{A_iY_{A_i}}{Y} = \frac{(A_iS_i)^q}{(A_1S_1)^q + (A_2S_2)^q}. \tag{15}$$

These expressions give (11). Since individual productivity $Z$ multiplies both returns, it scales output without changing the optimal weights. Maximised output is homogeneous of degree one and convex in skills: Workers can put more weight on whichever skill is stronger.

**Optimal investment.** Let $\mathcal{M}_i = \beta\mathbb{E}_t[V_{S_i}(S',B',Z')]$ be the marginal continuation value of skill $i$, and let $k_i = \rho I_i^{\rho-1}$ be its marginal training cost. Under the normalised costs $K_\rho$, this becomes $k_i = \kappa_\rho\rho I_i^{\rho-1}$. At a differentiable optimum with positive consumption,

$$\begin{aligned} u'(C)k_i - \mathcal{M}_i \geq 0, \qquad & I_i\left[u'(C)k_i - \mathcal{M}_i\right] = 0, \\ V_{S_i}(S,B,Z) = u'(C)Y_{S_i} + \Delta_i\mathcal{M}_i. & \end{aligned} \tag{16}$$

For positive current investment, substituting the envelope condition at the next date gives

$$u'(C)k_i = \beta\mathbb{E}_t\left[u'(C')Y'_{S_i} + \Delta_i\mathcal{M}'_i\right]. \tag{17}$$

If investment is also positive in every relevant next-period state, then $\mathcal{M}'_i = u'(C')k'_i$. Dividing by current marginal utility gives the Euler equation in (9). At a future investment corner, we instead retain the continuation value $\mathcal{M}'_i$ in (17).

The equation values an extra skill through both the output it produces and the future training it replaces. Payoffs in states with lower consumption receive greater weight when $\gamma > 0$. This does not give a general sign for the response of total investment to return uncertainty: A change in risk affects the payoffs, their association with consumption and future investment decisions. These are necessary conditions at differentiable optima. The

numerical solution searches jointly over both future skill stocks, rather than using an Euler root to select the optimum.

**Reference investment for the cost comparisons.** We use a deterministic symmetric steady state to define the reference investment $\bar{I}$ at which marginal training costs are matched. At constant effective returns, a positive interior steady state has $\bar{I}_i = (1-\Delta_i)\bar{S}_i$. Consumption is constant, so marginal utility cancels from the Euler equation:

$$(1-\beta\Delta_i)\rho\big[(1-\Delta_i)\bar{S}_i\big]^{\rho-1} = \beta Y_{S_i}(\bar{S},\bar{A}). \tag{18}$$

At unit returns and equal retention $\Delta$, the symmetric candidate is

$$\bar{S}_1 = \bar{S}_2 = \left[\frac{\beta\, 2^{1/q-1}}{\rho(1-\beta\Delta)(1-\Delta)^{\rho-1}}\right]^{1/(\rho-1)}. \tag{19}$$

The baseline parameters give $\bar{S}_i = 56.17$ and $\bar{I} = (1-\Delta)\bar{S}_i \simeq 2.81$. This is a reference from the interior steady-state conditions; those conditions alone do not establish global optimality, stability or uniqueness. Matching marginal cost at $\bar{I}$ leaves these conditions unchanged across the cost specifications. It does not match expenditure there or hold the desired skill portfolio fixed after returns change.

## C.2 Training Expenditure

Consider two periods, with current resources $W > 0$, CRRA coefficient $\gamma \geq 0$ and training-cost exponent $\rho > 1$. Skills depreciate fully, and the worker consumes all final-period output. Write $K = I_1^\rho + I_2^\rho$ and $p = I_1^\rho/K$. If $\lambda > 0$ multiplies both future skill returns, terminal output is

$$Y' = \lambda K^{1/\rho}\mathcal{Y}(p,A'), \qquad \mathcal{Y}(p,A') = \big[(A_1')^q p^{q/\rho} + (A_2')^q(1-p)^{q/\rho}\big]^{1/q}. \tag{20}$$

Returns are positive. We assume that the relevant expectations are finite and that an optimal allocation exists: $\mathbb{E}[\mathcal{Y}^{1-\gamma}]$ must be finite and positive for $\gamma \neq 1$, and $\mathbb{E}[|\log\mathcal{Y}|]$ must be finite for log utility. Positive lognormal returns satisfy these conditions. The derivatives below hold the distribution of $A'$ fixed and require $0 < K < W$.

CRRA utility separates the allocation of training from its scale. For every $K$ and $\lambda$, the optimal share maximises

$$\mathcal{H}_\gamma(p) = \begin{cases} \{\mathbb{E}[\mathcal{Y}(p,A')^{1-\gamma}]\}^{1/(1-\gamma)}, & \gamma \neq 1, \\ \exp\{\mathbb{E}[\log\mathcal{Y}(p,A')]\}, & \gamma = 1. \end{cases} \tag{21}$$

Hence, the optimal share does not depend on current resources, total expenditure or a

common scaling of future returns. For $\gamma > 1$, maximising this certainty equivalent means minimising the corresponding expected negative power of output. If $\rho > q$, output is strictly concave in the expenditures $I_i^\rho$ and has an infinite marginal benefit at either expenditure boundary, giving a unique interior allocation. The scale results need only an optimal allocation that is independent of scale.

Let $p^*$ be such an allocation and set $M_\gamma = \mathbb{E}[\mathcal{Y}(p^*, A')^{1-\gamma}]$ for $\gamma \neq 1$, with $M_1 = 1$. The expenditure condition is

$$(W - K)^{-\gamma} = \frac{\beta}{\rho}\lambda^{1-\gamma}K^{(1-\gamma)/\rho-1}M_\gamma. \tag{22}$$

At the optimal allocation, utility is strictly concave in $K$. For $\gamma > 0$, marginal benefits diverge at $K = 0$ and the marginal cost diverges at $K = W$, so the optimum is interior. Under risk neutrality, the unconstrained optimum is $K_u = [\beta\lambda\mathbb{E}[\mathcal{Y}(p^*, A')]/\rho]^{\rho/(\rho-1)}$. We require $K_u < W$ for the interior comparisons. If $K_u \geq W$, the problem with strictly positive current consumption has its supremum at the excluded boundary.

Taking logs of (22) gives the positive derivative

$$D = \frac{\gamma}{W - K} + \frac{\rho - 1 + \gamma}{\rho K} > 0. \tag{23}$$

Implicit differentiation then gives

$$\begin{aligned} \left.\frac{\partial K}{\partial \log \lambda}\right|_W &= \frac{1-\gamma}{D}, \\ \left.\frac{\partial K}{\partial W}\right|_\lambda &= \frac{\gamma/(W-K)}{D}. \end{aligned} \tag{24}$$

A common improvement in future returns therefore raises expenditure when $\gamma < 1$, leaves it unchanged when $\gamma = 1$, and reduces it when $\gamma > 1$. More current resources raise expenditure for $\gamma > 0$ and leave it unchanged at an interior risk-neutral optimum. Since the allocation stays fixed, both investment quantities have the same response sign as expenditure.

If current resources and both future returns rise together, with $W(\lambda) = \lambda\bar{W}$, then

$$\frac{dK}{d\log\lambda} = \frac{1 + \gamma K/(W-K)}{D} > 0. \tag{25}$$

This establishes the positive joint response for every $\gamma \geq 0$ at an interior optimum. It concerns a proportional change in both returns and current resources. An improvement in only one skill's return also changes the allocation, so this derivative does not sign that more general response.

**Log utility.** Full depreciation gives $S' = I$. Writing $I_1 = (Kp)^{1/\rho}$ and $I_2 = [K(1-p)]^{1/\rho}$ separates expenditure from its allocation. Since output is homogeneous of degree one in skills, final consumption is $C_1 = \lambda K^{1/\rho}\mathcal{Y}(p, A')$, while $C_0 = W - K$. Log utility therefore gives

$$\log(W - K) + \frac{\beta}{\rho}\log K + \beta \log \lambda + \beta \mathbb{E}[\log \mathcal{Y}(p, A')]. \tag{26}$$

The last two terms do not depend on $K$. The first-order condition for expenditure is

$$-\frac{1}{W-K} + \frac{\beta}{\rho K} = 0 \quad \Longleftrightarrow \quad \rho K = \beta(W - K).$$

The objective is strictly concave in $K$ on $(0, W)$, so rearranging this condition gives the unique optimal expenditure and current consumption:

$$K^* = \frac{\beta W}{\rho + \beta}, \qquad C_0^* = \frac{\rho W}{\rho + \beta}. \tag{27}$$

The rule holds for any admissible return distribution, including correlated or asymmetric returns. Changing that distribution can change the allocation and the skill quantities purchased, but leaves expenditure and current consumption unchanged at fixed $W$.

**Continued training under full depreciation.** This independence from future returns also holds with infinite-horizon training under log utility. Assume $\Delta_1 = \Delta_2 = 0$, $0 < \beta < 1$, exogenous returns, no additional scale-dependent constraints, and a well-defined finite value. This result concerns the continuous model, without fixed skill-grid bounds.

Take a feasible policy path from initial stocks $S$ and hold the return history fixed. Starting instead from $rS$, multiply date-$t$ stocks, output, expenditure and consumption by $r^{\rho^{-t}}$, and investment quantities by $r^{\rho^{-(t+1)}}$. Full depreciation and the degree-$\rho$ cost function preserve the stock laws and budgets. Discounted log utility changes by

$$\log r \sum_{t=0}^{\infty} (\beta/\rho)^t = \frac{\rho}{\rho - \beta}\log r.$$

This change is independent of the policy, and the construction is reversible. Hence,

$$V(rS, B, Z) = V(S, B, Z) + a \log r, \qquad a = \frac{\rho}{\rho - \beta}. \tag{28}$$

The expenditure part of the Bellman problem is therefore $\log(W - K) + (\beta a/\rho)\log K$, giving

$$K^* = \frac{\beta a}{\rho + \beta a} W = \frac{\beta}{\rho} W, \qquad C^* = \left(1 - \frac{\beta}{\rho}\right) W. \tag{29}$$

Continued training alone does not remove the spending invariance. Retained skills do

invalidate this proof: Scaling current resources by $r$ scales retained stocks by $r$ but newly purchased skills by $r^{1/\rho}$. Their sum no longer has the required scaling. This does not determine the sign of the expenditure response to uncertainty in the full model.

## C.3 Uncertainty and Skill Diversification

This subsection proves Proposition 1. Fix current resources $W > 0$, retained skills $d_i \geq 0$, $\beta \in (0,1)$ and $\rho > q > 1$. The worker chooses $I_i \geq 0$, with $K = I_1^\rho + I_2^\rho < W$, to maximise $\log(W-K) + \beta \mathbb{E}[\log Y(d+I, A')]$. All final output is consumed, with no further training. Let $T_i = d_i + I_i$ denote final stocks.

We initially allow arbitrary positive expected returns $\bar{A}_i$. Returns follow

$$\log \frac{A_i'}{\bar{A}_i} = -\frac{v_C}{2} - \frac{v}{8} + \xi_C + (-1)^{i+1}\frac{\xi_R}{2}, \qquad \xi_C \sim N(0, v_C), \quad \xi_R \sim N(0, v), \tag{30}$$

with independent shocks. Here $v = v_R$ is log relative-return variance, and $v_C$ is held fixed. The mean corrections keep $\mathbb{E}[A_i'] = \bar{A}_i$. The common component adds a constant to expected log output and therefore does not affect investment.

**A unique interior optimum.** Write expenditure as $z_i = I_i^\rho$ and set $a = 1/\rho$. Each term $f_i(z_i) = (d_i + z_i^a)^q$ in the inner production sum satisfies

$$f_i''(z_i) = qaz_i^{a-2}(d_i + z_i^a)^{q-2}\left[(a-1)d_i + (aq-1)z_i^a\right] < 0. \tag{31}$$

Hence the Hessian of terminal log output, $q^{-1}\log\sum_i (A_i')^q f_i(z_i)$, is negative definite: Its diagonal term is strictly negative and its outer-product term is negative semidefinite. Taking expectations and adding $\log(W - z_1 - z_2)$ preserves strict concavity.

The feasible expenditure closure is compact. Expected utility is finite at positive stocks and positive current consumption, and tends to minus infinity when current consumption vanishes or both final stocks vanish. It therefore attains a maximum away from those boundaries. If either expenditure is zero while the other final stock is positive, its marginal terminal benefit is infinite. Thus both investments are positive. Strict concavity makes this optimum unique.

**The response to relative-return risk.** Define $h = (\log T_1 + \log T_2)/2$, $x = \log(T_1/T_2)$ and $\ell = \log(\bar{A}_1/\bar{A}_2)$. Up to terms independent of $(h, x)$, the objective is

$$F(h, x; v) = \log\left[W - \sum_{i=1}^{2}\left(e^{h+(-1)^{i+1}x/2} - d_i\right)^{\rho}\right] + \beta h + \beta H(x, v),$$
$$H(x, v) = \frac{1}{q}\mathbb{E}\left[\log\left(2\cosh\frac{q(\ell + x + \sqrt{v}\,N)}{2}\right)\right], \qquad N \sim N(0, 1). \tag{32}$$

Since $H$ does not depend on $h$, the locally optimal scale conditional on $x$, denoted $h^*(x)$, does not depend on $v$. Let $\Phi(x, v) = F(h^*(x), x; v)$. At the optimum,

$$\Phi_{xx} = F_{xx} - \frac{F_{xh}^2}{F_{hh}} < 0. \tag{33}$$

The inequality follows because the smooth invertible change from expenditures to $(h, x)$ preserves negative definiteness of the Hessian at a stationary point.

For $g(u) = (1 + e^{-u})^{-1}$, Gaussian differentiation gives

$$H_x = \mathbb{E}[g(q(\ell + x + \sqrt{v}\,N))] - \frac{1}{2},$$
$$H_{xv} = \frac{q^2}{2}\mathbb{E}[g''(q(\ell + x + \sqrt{v}\,N))]. \tag{34}$$

The function $g''$ is odd and negative for positive arguments. A normal variable $U$ with positive mean has density $f_U(u) > f_U(-u)$ for $u > 0$, so

$$\mathbb{E}[g''(U)] = \int_0^{\infty} g''(u)[f_U(u) - f_U(-u)]\,du < 0.$$

The sign reverses at a negative mean and is zero at a zero mean. At $v = 0$, the same result follows from the right derivative of the Gaussian expectation. Interiority and the nonsingular Hessian allow implicit differentiation at every finite variance, giving

$$\frac{dx^*}{dv} = -\frac{\beta H_{xv}}{\Phi_{xx}}, \qquad \operatorname{sgn}\left(\frac{dx^*}{dv}\right) = -\operatorname{sgn}(\ell + x^*). \tag{35}$$

Equal expected returns imply $\ell = 0$, proving the proposition at every finite $v \geq 0$. Since $\operatorname{sp}(T)^2 = 1 - 1/\cosh x$, reducing $|x|$ also reduces specialisation.

To identify the investment changes, use the scale condition

$$W = \sum_{i=1}^{2}\left[I_i^{\rho} + \frac{\rho}{\beta}T_i I_i^{\rho-1}\right]. \tag{36}$$

Each summand increases strictly in its own $T_i$, holding $d_i$ fixed. Along this condition, a fall

in $x$ requires lower $T_1$ and higher $T_2$. Hence, at equal expected returns, greater uncertainty reduces investment in the skill with the larger chosen final stock and increases investment in the other. Investment itself need not become more evenly divided because inherited stocks may differ. Neither total investment nor expenditure must remain constant.

With unequal expected returns, the response instead moves the ratio of $\bar{A}_i T_i$ towards balance and can increase specialisation in raw stocks. The proof establishes the direction of adjustment, not convergence to equal stocks as variance grows. Continued training and persistent realised return differences are outside this two-period result.

## C.4 Return Processes and Numerical Solution

The original return process has persistent common and individual components. Their centred logs follow

$$\begin{aligned} b_{i,t+1} &= \rho_A b_{i,t} + \sqrt{1-\rho_A^2}\, s_B \eta_{i,t+1}, \\ z_{n,t+1} &= \rho_A z_{n,t} + \sqrt{1-\rho_A^2}\, s_Z \varepsilon_{n,t+1}, \end{aligned} \tag{37}$$

with independent standard normal innovations across dates and components. Individual innovations are also independent across workers, while all workers face the same $b_i$. Levels are $B_i = \exp(b_i - s_B^2/2)$ and $Z_n = \exp(z_n - s_Z^2/2)$, so both components have stationary mean one. Conditional expected returns still depend on the current state.

Table C1 gives the parameterisation. We split the stationary log variance of each effective return equally between its common and individual components. This gives variance $0.95^2$ and correlation $1/2$ between a worker's two effective log returns, since the same $Z_n$ multiplies both skills. These choices are illustrative; we do not fit them to the local projections or equate a model period with a survey wave.

Table C1: Illustrative model parameters

| Parameter | Interpretation | Value |
|---|---|---|
| $\beta$ | Discount factor | 0.99 |
| $\gamma$ | CRRA coefficient | 1 |
| $\rho$ | Training-cost exponent | 2.5 |
| $\sigma$ | Skill-use constraint curvature | 2 |
| $\Delta_1, \Delta_2$ | Skill retention rates | 0.95 |
| $\rho_A$ | Log-component persistence | 0.9 |
| $s_B, s_Z$ | Stationary log standard deviations | $0.95/\sqrt{2}$ |

The baseline grid has 600 equally spaced points for each skill on $[0.01, 200]$. We approximate each return component by a three-state Rouwenhorst chain, giving 27 joint return

states, and normalise each exponentiated component grid to have stationary mean one. We solve the Bellman problem by Howard policy iteration, searching jointly over feasible future skill stocks. Investment must be non-negative and consumption positive. The baseline Bellman residual is $9.75 \times 10^{-11}$. This checks the finite-grid solution, without removing approximation error from discretising skills and returns.

Common returns make the cross-sectional distribution depend on the common history. Conditional on that history, we propagate the distribution over $(S_1, S_2, Z)$ using investment policies and individual transitions. A distribution that also averages over common histories is therefore not the cross-section at a particular date. All reported dynamic expectations integrate over future common and individual histories on the finite grid. They have no Monte Carlo sampling error.

## C.5 Return and Uncertainty Experiments

To change uncertainty while holding expected return paths fixed, we use a separate reference process for the common returns. It retains their original three-point support $b$ and stationary probabilities $\pi$, but replaces their transition matrix by

$$P^0 = r_B \mathbb{I} + (1 - r_B)\mathbf{1}\pi, \qquad r_B = 0.9. \tag{38}$$

This retains the current return with probability $r_B$ and otherwise redraws from $\pi$. Here persistence concerns levels, whereas $\rho_A$ concerns logs. Individual productivity keeps its original process.

The higher-risk matrix $P^H$ spreads each row of $P^0$ while preserving its mean. On support $x_0 < x_1 < x_2$, we move probability $d$ away from the middle state:

$$h - q = d\left(\frac{x_2 - x_1}{x_2 - x_0}, -1, \frac{x_1 - x_0}{x_2 - x_0}\right). \tag{39}$$

The variance increase is $d(x_1 - x_0)(x_2 - x_1)$. We choose $d$ in each row to raise its variance by 20%, with non-negative resulting probabilities. Since the conditional mean remains affine in the current return, both laws satisfy

$$\mathbb{E}_r[B_{t+h} \mid B_t = b_j] = \bar{b} + r_B^h(b_j - \bar{b}), \qquad \bar{b} = \pi b, \quad r \in \{0, H\}. \tag{40}$$

Thus expected level returns agree at every horizon, not only at the next draw. The two common skill returns remain independent. The experiments with a 10% variance increase use the same construction.

We solve separately for the investment policies under $P^0$ and $P^H$, then evaluate both under $P^0$. This holds realised return histories fixed while changing the uncertainty un-

der which workers choose their investment. The higher-risk policy is optimal for $P^H$; evaluating it under $P^0$ does not describe an additional rational-expectations equilibrium under that law. For the uncertainty-only and temporary-return experiments, all training-cost specifications start from the same reference invariant distribution, conditional on middle common returns. Each specification subtracts its own evolving path without the intervention.

**Temporary return improvements.** We add $b^* = 1.1b_{\text{mid}}$ to the skill-1 support and solve the worker's problem at this state. The reference process retains $b^*$ with probability $r_B$ and otherwise redraws from the original support. Original states cannot enter the added state, so their policies are unchanged. Hence,

$$\mathbb{E}[B_{1,t} \mid B_{1,0} = b^*] - \mathbb{E}[B_{1,t} \mid B_{1,0} = b_{\text{mid}}] = r_B^t(b^* - b_{\text{mid}}). \tag{41}$$

The return advantage therefore declines by 10% each period in expectation. The extended grid has 36 joint return states.

**Permanent return improvements.** The baseline permanent-return exercise multiplies $B_1$ by 1.1 along every history of the original process, with unchanged transition probabilities. Its initial population follows a common-history burn-in, conditional on middle common returns. The comparison with greater uncertainty instead uses $P^0$ and $P^H$, keeping these same initial workers and evaluating both policies under $P^0$. It therefore recomputes the permanent-return response under the matched process. These are different experiments, so their numerical gains need not coincide.

## C.6 Training and Output Accounting

**Current resources and future payoffs.** At given skills and individual productivity, let $a \in \{0, 1\}$ choose current resources $W_a$ before or after the return improvement, and let $b \in \{0, 1\}$ choose the corresponding conditional distribution of future returns. We solve

$$\max_{I \geq 0,\, K(I) < W_a} u\big(W_a - K(I)\big) + \beta \mathbb{E}_b[V(\Delta S + I, B', Z')]. \tag{42}$$

All four choices use the same continuation-value function on the extended return support. We change current resources and the predictive content of current returns separately, without changing the subsequent transition law.

For any resulting outcome $X_{ab}$, we average each contribution over the two orders of

introducing the changes:

$$\begin{aligned} \bar{R}_X &= \tfrac{1}{2}[(X_{10} - X_{00}) + (X_{11} - X_{01})], \\ \bar{P}_X &= \tfrac{1}{2}[(X_{01} - X_{00}) + (X_{11} - X_{10})], \\ X_{11} - X_{00} &= \bar{R}_X + \bar{P}_X. \end{aligned} \tag{43}$$

The future-payoff contribution changes the entire conditional distribution, including means and risk. It therefore does not isolate a change in expected returns alone. These are comparisons within the model, not empirical mediation estimates.

**Sources of the output gain.** For a paired shocked and reference history, let $B_t^1, B_t^0$ be common returns and $\mu_t^1, \mu_t^0$ the conditional distributions of skills and individual productivity. Write $y(S, B, Z; J) = Z \sum_i B_i J_i S_i$, let $y^*$ maximise this output over feasible weights, and set $J_t^0(S) = J^*(S, B_t^0)$. We change returns, skill use and skill stocks in that order:

$$\begin{aligned} D_t^{\text{return}} &= \int \left[ y(S, B_t^1, Z; J_t^0(S)) - y^*(S, B_t^0, Z) \right] d\mu_t^0, \\ D_t^{\text{use}} &= \int \left[ y^*(S, B_t^1, Z) - y(S, B_t^1, Z; J_t^0(S)) \right] d\mu_t^0, \\ D_t^{\text{skills}} &= \int y^*(S, B_t^1, Z)\, d\mu_t^1 - \int y^*(S, B_t^1, Z)\, d\mu_t^0, \\ \Delta Y_t &= D_t^{\text{return}} + D_t^{\text{use}} + D_t^{\text{skills}}. \end{aligned} \tag{44}$$

The skill-use contribution is non-negative because the reference weights remain feasible. The skills contribution includes changes in their total amount, composition and association with individual productivity, as well as interactions with improved returns.

For the temporary improvement, the two paths share retention and redraw events and receive the same return at their first redraw. Their returns then coincide, although earlier investment can leave different skill stocks. The identity holds for each paired history and after averaging. Skills contribute zero at impact. Around the original optimum, skill-use adjustment contributes at second order in the return change, while the direct effect is first order.

Figure C1 reports these contributions. Accumulated skills account for about 33% of the remaining output gain after ten periods and 60% after twenty, as the direct return advantage fades.

The contributions depend on their stated order. Permanently restricting an adjustment margin would change investment decisions as well, so this accounting does not give welfare

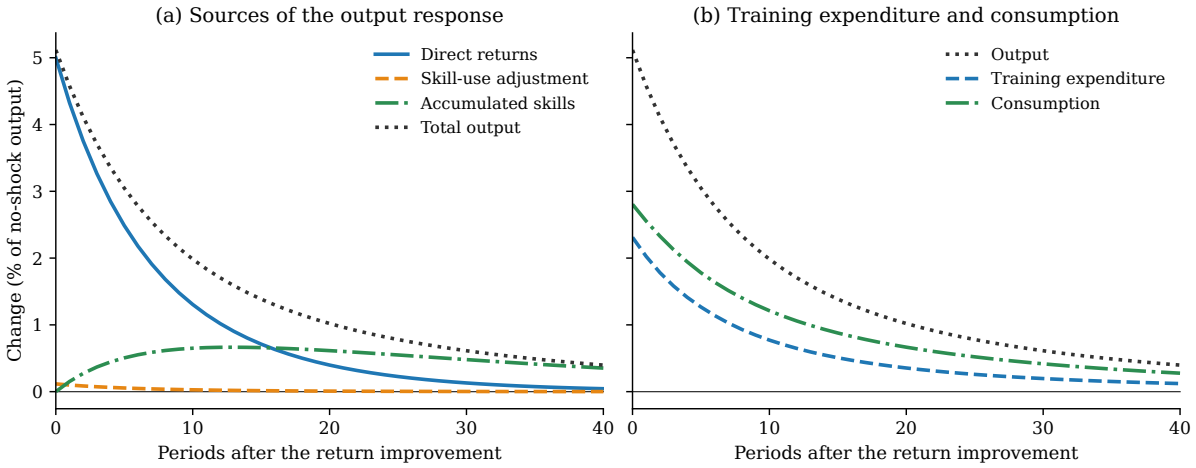



Figure C1: Worker adjustment and the output response

*Notes:* Contributions after the temporary 10% increase in $B_1$, as percentages of reference output. Panel (a) applies (44); panel (b) divides additional output between training expenditure and consumption. Initial workers and return histories match Figure 9. Reference and shocked skill-1 returns share retention and redraw events. Horizons are model periods.

effects. The resource and wage identities are

$$\Delta C_t = \Delta Y_t - \Delta K_t, \qquad \Delta w_t = (1-\theta)\Delta Y_t + \theta \Delta C_t. \tag{45}$$

Thus the level change in wages lies between the level changes in output and consumption for a common fixed training-payment share.

**Training expenditure and skill composition.** More training quantity can coexist with lower expenditure because dividing investment more evenly reduces its cost. For the baseline $\rho = 2.5$ and $\kappa_\rho = 1$, write $T = I_1 + I_2$ and $a_\rho = 2^{1-\rho}$. Then

$$\begin{aligned}\mathbb{E}[K] = a_\rho \big(\mathbb{E}[T]\big)^\rho + a_\rho \left\{\mathbb{E}[T^\rho] - \big(\mathbb{E}[T]\big)^\rho\right\} \\ + \mathbb{E}\left[I_1^\rho + I_2^\rho - a_\rho T^\rho\right].\end{aligned} \tag{46}$$

The terms capture the cost of mean quantity split equally, dispersion in total quantity, and unequal allocation within each portfolio. At period one hundred, their changes contribute 0.472, 0.277 and $-1.206$ percentage points of reference expenditure, respectively. Hence expenditure falls by 0.457% despite greater total quantity. This is an identity for the resulting distributions, rather than a comparison holding every worker's quantity fixed.

To separate skill quantities and composition in output, condition on realised returns $a = (B_1, B_2, Z)$. Let $L = S_1 + S_2$, $m_r(a) = \mathbb{E}_r[L \mid a]$ and $v_r(a) = \mathbb{E}_r[Y \mid a]/m_r(a)$ under policy $r \in \{0, H\}$. Since realised histories are the same, both policies have return-state

probabilities $p(a)$. With bars denoting the average across policies,

$$\Delta\mathbb{E}[Y] = \sum_a p(a)\,\Delta m(a)\,\bar{v}(a) + \sum_a p(a)\,\bar{m}(a)\,\Delta v(a). \tag{47}$$

The first term changes skill quantities across return states; the second changes skill-weighted portfolio composition within states. At period one hundred they contribute 0.303 and $-0.598$ percentage points of reference output. Composition here includes which skill pays more, rather than only the unsigned specialisation index.

Allowing realised returns to follow $P^H$ changes the comparison. Long-run mean specialisation then rises by 0.443 percentage points. Changing return-state probabilities contributes 1.077 points, while changing conditional mean specialisation contributes $-0.634$ points. Conditional specialisation falls in 23 of the 27 states, but the different distribution of realised opportunities reverses the aggregate response. These conditional distributions also reflect earlier histories. The diversification result under unchanged histories therefore does not imply that greater realised return risk lowers unconditional specialisation.

## C.7 Numerical Sensitivity

For Table 2, we bound current choices over the feasible interpolation rectangles of the same bilinear continuation value. The expenditure response and both contributions retain their reported signs throughout the bounds. The bounds include current optimisation and finite-grid Bellman errors, but exclude interpolation and skill- or return-grid approximation errors relative to the continuous-state model.

Repeating the temporary-return experiments with 400 rather than 600 skill points gives similar output responses. Specialisation magnitudes are more sensitive. At period twenty, the two specialised workers' responses are 0.16 and $-0.80$ percentage points on the finer grid, compared with 0.56 and $-0.47$ on the coarser grid. Their opposite directions at that horizon survive, while some earlier signs change. The uncertainty-only experiment also preserves the ordering across curvatures: Period-twenty output responses at $\rho = 2.5$ and 4 round to $-0.30\%$ and $-0.09\%$ on both grids.

The accounting identities, probability mass and resource constraints hold to numerical precision in the reported calculations. These checks establish internal consistency on the chosen grids; they do not establish continuous-state convergence or uniqueness of the invariant distribution.