Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data*

Leonid Kogan
MIT Sloan School of Management and NBER

Dimitris Papanikolaou
Kellogg School of Management and NBER

Lawrence D. W. Schmidt
MIT Sloan School of Management

Bryan Seegmiller
Kellogg School of Management

October 2023

Abstract

We develop measures of labor-saving and labor-augmenting technology exposure using textual analysis of patents and job tasks. Using US administrative data, we show that both measures negatively predict earnings growth of individual incumbent workers. While labor-saving technologies predict earnings declines and higher likelihood of job loss for all workers, labor-augmenting technologies primarily predict losses for older or highly-paid workers. However, we find positive effects of labor-augmenting technologies on occupation-level employment and wage bills. A model featuring labor-saving and labor-augmenting technologies with vintage-specific human capital quantitatively matches these patterns. We extend our analysis to predict the effect of AI on earnings.

*We are grateful to Daron Acemoglu, David Autor, Effi Benmelech, Nicholas Bloom, Carter Braxton, Julieta Caunedo, Martin Beraja, Carola Frydman, Tarek Hassan, Anders Humlum, Michael Peters, Pascual Restrepo, Jonathan Rothbaum, Miao Ben Zhang, along with seminar participants at University of Amsterdam, BI-SHoF Conference, Boston University, CIREQ Macroeconomics Conference, Columbia GSB, FIRS, Johns Hopkins, HKUST, Labor and Finance Group, NBER (EFG, PRMP), Macro-Finance Society, MIT Sloan, Michigan State, the Econometric Society, Rice University, University of Rochester, the Society of Economic Dynamics, University College London, University of Illinois at Urbana Champaign, University of Toronto, Tsinghua PBC, WFA, and Wharton for valuable discussions and feedback. We thank Carter Braxton, Will Cong, and Jonathan Rothbaum for generously sharing code. Huben Liu provided outstanding research support. The paper has been previously circulated under the titles "Technological Change and Occupations over the Long Run", "Technology-Skill Complementarity and Labor Displacement: Evidence from Linking Two Centuries of Patents with Occupations," and "Technology, Vintage-Specific Human Capital, and Labor Displacement: Evidence from Linking Patents with Occupations". The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7503840, Disclosure Review Board (DRB) approval numbers: CBDRB-FY21-POP001-0176, CBDRB-FY22-SEHSD003-006, CBDRB-FY22-SEHSD003-023, CBDRB-FY22-SEHSD003-028,CBDRB-FY23-SEHSD003-0350, CBDRB-FY23-SEHSD003-064).

Economists and workers alike have long worried about the prospect of technological displacement of labor. New technologies can displace incumbent workers, either because these technologies can now perform certain tasks formerly done by workers, or because these technologies require new skills that incumbent workers lack. Labor-saving technologies directly substitute for labor and can lead to a reduction in wages as demand for labor falls. By contrast, labor-augmenting technologies can benefit workers as a whole, even though individual workers who cannot adapt to the new technologies may be left behind. The distinction between these two channels is relevant for understanding the effects of technology on the cross-section of workers: automation of tasks performed by a certain occupation displaces all workers if wages fall, whereas the displacive effects of the development a new labor-augmenting method of production will be concentrated to the workers most skilled in the previous technology that lack the ability to adapt. We provide evidence that is consistent with both of these two channels, by first developing direct measures of workers' exposures to these two types of technologies, and then linking it to individual worker outcomes using administrative data.

We begin by developing new direct measures of workers' exposure to technological innovation using textual analysis of patents and occupation task descriptions. Our measures vary both in the time series and in the cross-section, and, importantly, differentiate between labor-saving and labor-augmenting technologies. Labor-saving technologies (automation) can perform routine tasks in place of human workers. Our assumption is that these technologies are related to tasks performed by workers that can be classified as routine (or alternatively, require little related experience). Labor-augmenting technologies, by contrast, complement certain worker tasks, and therefore are related to tasks that can be classified as non-routine (or alternatively, require high levels of related experience). We then link our exposure measures to individual worker outcomes: using administrative data on worker earnings from the United States, we show that both types of technologies can displace incumbent workers through different mechanisms—direct substitution versus skill displacement. New labor-saving technologies displace workers regardless of their characteristics, whereas new labor-augmenting technologies displace the most skilled and older incumbent workers in favor of new entrants.

These two notions of technology exposure arise naturally in a model in which different occupations perform a mix of different routine and non-routine tasks. To formalize this intuition, we develop a model that combines elements of Violante (2002), Caunedo, Jaume, and Keller (2023), and Acemoglu and Restrepo (2018). We model technological innovation as declines in the quality-adjusted price of capital that is specific to each task. Different innovations can simultaneously affect different tasks. The key difference between these two types of tasks is that 1) the elasticity of substitution between capital and labor is greater in routine than in non-routine tasks; and 2) technological improvements in capital that complements labor in production entails some degree of skill displacement.

In our model, a given worker's exposure to labor-saving or labor-augmenting technologies can vary over time. Occupations vary in their exposure to these two types of technologies because workers often perform a mix of tasks, and a single technological advancement can enhance the quality of capital used across different routine and non-routine tasks. Occupations with high exposure to automation technologies at a specific point in time are those occupations performing a higher share of routine tasks that experienced significant capital improvements during that period. Similarly, occupations most exposed to new labor-augmenting technologies perform a higher share of non-routine tasks for which capital improved during this period. Since the arrival rate of new technologies is neither constant nor uniform across tasks, different occupations are exposed to technology improvements at different points in time, even if their share of routine tasks is constant.

We use textual analysis of patents and occupation task descriptions to construct technology exposure measures that closely resemble their counterparts generated by the model. First, we classify each task performed by an occupation as a routine or non-routine (alternatively as high- or low-required experience) using generative AI (GPT4). The resulting task classifications aggregated at the occupation level correlate highly with the Acemoglu and Autor (2011) measure of routine-task-intensity (RTI) and ONET's classification of jobs into high and low-skill, respectively. Second, we use natural language processing methods to estimate the similarity between the textual description of the routine or non-routine tasks performed by an occupation and that of major technological breakthroughs. We identify these breakthroughs using the methodology of Kelly, Papanikolaou, Seru, and Taddy (2021) who define a breakthrough innovation as one that is both novel (i.e. distinct from prior patents) and impactful (i.e. related to subsequent patents). By exploiting the timing of patent grants, we measure an individual worker's exposure to both labor-saving and labor-augmenting technologies at a given point in time.

We next examine how our measures of technology exposure relate to individual worker outcomes. We rely on employer-employee matched administrative earnings records from the US Social Security Administration, starting in the early 1980s, which are linked with information on each worker's occupation and education from the Current Population Survey. Thus, relative to the literature which has mostly studied repeated cross-sections of workers, we are able to measure a worker's occupation prior to the introduction of related technologies, and estimate how her earnings evolve in future years even if she switches employers, industries, and/or occupations.

Our empirical analysis leverages the granularity of our patent-occupation measures to generate time-variation in worker exposure at the industry-occupation level: this variation is driven by differences in the rate at which firms in different industries develop and patent new technologies that are related to the routine and non-routine tasks of a given occupation at a point in time. Our empirical estimates are identified by comparing two workers in different industry–occupation cells that are otherwise similar in their observable characteristics and past earnings history. Initially, we focus on the direct effect of technology exposure, by controlling for variation related to changes in industry productivity—by including industry-year together with occupation-year and worker income rank-year fixed effects in our specification. Our specification allows us to control for common shocks to labor demand and supply at the industry or occupation levels.

We find that improvements in labor-saving (automation) technologies are negatively related to the wage earnings of workers in affected occupation–industry cells. For instance, an increase in our exposure measure from the median to the 90th percentile is associated with a 2.5 percentage point decline in the total earnings of the average worker over the next five years. These earnings losses are concentrated on a subset of workers, since exposed workers experience a 1.2 percentage point increase in the probability of involuntary job loss over the next five years. Since we do not observe layoffs, we use as a proxy cases in which a worker switching employers experiences a change in earnings in the bottom 20th percentile. These effects are pervasive for both blue- and white-collar workers: our point estimates are similar across workers in both manufacturing and services, as well as in occupations that emphasize manual or cognitive skills. Importantly, the magnitude of these wage declines or job loss probabilities are essentially unrelated to observable measures of worker skill—measured by age, level of wage earnings relative to other workers in the same industry and occupation, and college education.

Perhaps surprisingly, but consistent with our model, new labor-augmenting technologies also lead to a decline in earnings for exposed workers, though the average magnitudes are smaller. An increase in our exposure measure from the median to the 90th percentile is associated with a 1.3 percentage point decline in earnings growth and a 0.5 percentage point increase in the likelihood of involuntary job loss. However, unlike in the case of labor-saving technology, the effects of exposure to labor-augmenting technologies are fairly heterogenous: it disproportionately affects white-collar workers (defined as those with college degree, or those employed in non-manufacturing industries or in occupations emphasizing cognitive tasks); older workers; and workers that are paid more relative to their peers (other workers with similar characteristics in the same industry and occupation). For example, a worker in the top 5 percent of pay relative to her peers experiences a 4.2 percentage point decline in her average earnings over the next five years and a 3.8 percentage point increase in the likelihood of involuntary job loss. The estimated coefficients are weakly negative but not statistically different from zero for workers below the 50th percentile in relative pay.

The empirical findings discussed so far capture the direct effects of technology on incumbent workers. In the model, these direct effects operate through changes in the price of labor resulting from technology improvements, or skill displacement, while holding total demand for labor fixed. To study the broader effects of technology on the labor market, we also trace the impact of technology on new entrants and spillovers to all workers. These spillovers occur because new technologies also affect the overall demand for labor even for workers whose tasks are not affected by technological improvements.

Our model implies that labor-saving and labor-augmenting technologies should both increase productivity, but that these two types of improvements can have different effects on employment and the labor share. Using aggregate data at either the occupation–industry or industry level, we find that both exposure measures are positively related to industry productivity. However, exposure to labor-saving technologies is associated with lower employment and a decline in the labor share, while exposure to labor-augmenting technologies is associated with a significant increase in employment and wages, and a weakly higher labor share. This sharp difference between the response of wages and employment to labor-augmenting technologies estimated from individual versus the aggregate data is consistent with the redistributive effects of skill displacement, in which new vintages of labor-augmenting technologies primarily benefit new workers. Last, we find some evidence consistent with spillovers: the average technology exposure of an industry is associated with increased earnings for workers after controlling for their own direct exposure. Our estimates based on the micro and aggregate data are quantitatively consistent with the model calibrated to a common set of parameters that are in line with the literature.

In the last part of our paper, we apply our findings to the case of Artificial Intelligence (AI). As a general-purpose technology, AI has both labor-augmenting and labor-saving applications that can potentially affect a large number of occupations. We first construct an estimate of the potential for AI to substitute or complement the tasks performed by each occupation; we then leverage our estimates from the calibrated model to generate predictions for the likely growth path of wage earnings for incumbent workers in each occupation. While these estimates are highly speculative, we estimate that AI is likely to primarily displace white-collar workers in office and administrative occupations, and to some extent workers in production and transportation. That said, AI has the potential to significantly increase the earnings of certain occupations that can leverage AI to increase their productivity (for example, personal financial advisors), assuming the skill displacement effects are small.

Overall, our work contributes to the existing literature along two key dimensions. First, we propose (and validate) a new direct measure of workers' technology exposure that differentiates between labor-saving and labor-augmenting technologies and varies over time. Second, we provide a systematic analysis of the heterogeneous effects of both labor-substituting and labor-augmenting technologies on worker-level earnings outcomes in the United States since the early 1980s. Our focus on individual worker data allows us to elucidate two distinct channels through which improvements in technology can displace incumbent workers: 1) new technologies perform the same tasks workers do, but at a lower cost; 2) new technologies helps the workers do the job more efficiently, but using these technologies efficiently may require new skills that certain incumbent workers may lack. Considering both of these channels paints a more nuanced view of the effects of technology on the labor market.

Existing measures of workers' technology exposure have mainly focused on labor-saving technologies, with a particular emphasis on the adoption of robots.[1] While focusing on narrow measures of automation helps us understand the effects of particular technologies, they paint an incomplete picture of the impact of automation on workers: Benmelech and Zator (2022) argue that investment in robots is small and highly concentrated in a few industries, and accounts for less than 0.3% of aggregate expenditures on equipment. In comparison, our measure of automation is broader and identifies technologies that relate to both blue and white collar workers. Mann and Püttmann (2018); Dechezleprêtre, Hémous, Olsen, and Zanella (2021) provide somewhat broader measures of automation using patents; they use keyword-based classification methods to identify automation patents in more recent periods. Relative to these approaches, our measure allows us to link the tasks performed by specific occupations to these labor-saving technologies, which allows us to also explore cross-occupation variation in our empirical analysis. In this regard, our approach is similar to Webb (2020) who also analyzes the similarity between patents identified as being related to robots, AI, or software and occupation task descriptions. Relative to Webb (2020), our measure differentiates between technologies related to workers' routine or non-routine tasks which allows us to distinguish between labor-saving and labor-augmenting technologies, and varies over time which allows us to control for occupation or industry-specific time trends.

Focusing on individual workers provides a fuller picture of the relation between technology and worker earnings than focusing on average worker outcomes, which likely obscures important heterogeneity and is affected by composition effects. Existing work studying the impact of automation on individual workers has focused on European labor markets over short time periods.[2] The main focus thus far has been on the impact of firms' adoption of robots (Humlum, 2019), or automation technologies (Bessen, Goos, Salomons, and van den Berge, 2023). Other work exploits more aggregate sources of variation (Dauth et al., 2021). The evidence is somewhat mixed: Humlum (2019) and Bessen et al. (2023) find negative effects for production workers, especially older workers; Dauth et al. (2021) finds a positive effect on the earnings of incumbent workers but negative effects on new entrants. A key advantage of our approach is that we are able to link specific labor-saving technologies to specific workers in a given industry and occupation. Examining a broad range of labor-saving technologies over a longer period in the United States, we find negative effects on workers affected by these technologies, across both blue collar and white collar workers, with both

[1] An incomplete list includes Acemoglu and Restrepo (2020, 2022); Graetz and Michaels (2018); Humlum (2019); Dauth, Findeisen, Suedekum, and Woessner (2021); Koch, Manuylov, and Smolka (2021); Bonfiglioli, Crinò, Fadinger, and Gancia (2020).

[2] A notable exception is Feigenbaum and Gross (2020), who finds that women were more likely to be in lower-paying occupations following the adoption of mechanical switching technology by AT&T (a labor-saving technology) in the early 20th century in the United States.

the youngest and oldest workers experiencing similarly adverse effects.[3]

In contrast to automation technologies, the impact of labor-augmenting technologies on individual worker earnings has received significantly less attention: Akerman, Gaarder, and Mogstad (2015) investigates the adoption of broadband internet on earnings of Norwegian workers and finds evidence in favor of skill complementarity. Autor, Chin, Salomons, and Seegmiller (2022) develop a measure of labor-augmenting technology exposure based on Census job description write-ins, however, their focus is on analyzing the factors generating demand for new tasks and their implications for aggregate employment instead of individual workers. We contribute to this literature by documenting that labor-augmenting technologies are negatively related to the earnings of directly exposed incumbent workers—especially the oldest and most highly paid workers. This decline in the earnings of incumbents stands in sharp contrast with the increased earnings of all workers in the same industry and occupation, which suggests that new hires, or other workers in the same industry, reap the benefits of labor-augmenting technologies.

We interpret this negative exposure of more skilled workers as evidence for vintage-specific human capital (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). These models assume that workers who have accumulated greater levels of skill in existing technologies are more likely to be displaced when newer vintages of technology arrive. Thus, our work also complements the existing empirical evidence emphasizing skill displacement (Deming and Noray, 2020; Braxton and Taska, 2020; Hombert and Matray, 2021; Atack, Margo, and Rhode, 2019, 2022). Closest to our paper is Braxton and Taska (2020), who show that displaced workers in occupations facing changing skill requirements fare worse than other displaced workers; Deming and Noray (2020), who finds that the wage premium of jobs with frequent changes in skill demands (STEM) declines with worker age; and Hombert and Matray (2021) who show that French workers starting out in the ICT sector right at the dot com crash had large declines in wages relative to other sectors.

More generally, our work is connected to the literature aiming to understand the complementarity between capital (broadly defined) and labor. Existing work emphasizes the complementarity between technology and certain types of worker tasks or skills (Goldin and Katz, 1998, 2008; Autor, Levy, and Murnane, 2003; Autor, Katz, and Kearney, 2006; Goos and Manning, 2007; Autor and Dorn, 2013; Autor et al., 2022); or the substitution between workers and capital (Krusell, Ohanian, Ríos-Rull, and Violante, 2000; Hornstein, Krusell, and Violante, 2005, 2007; Karabarbounis and Neiman, 2013; Hemous and Olsen, 2021; Caunedo et al., 2023). The main focus of these papers is how technology

[3] To reconcile our estimates with existing studies based on European data, one possibility is that institutional differences across U.S. and European labor markets may lead to different effects on worker earnings. Studies of automation based on aggregate data also reveal sharp differences between these two regions. For example, Acemoglu and Restrepo (2020, 2022) argue that automation has resulted in large negative demand shifts in the United States. Instead, Graetz and Michaels (2018) find a positive impact of automation technologies on productivity across a sample of 17 countries, with little evidence of a negative effect on local labor demand, while Aghion, Antonin, Bunel, and Jaravel (2021) reports similar findings at the firm level in France.

affects differences in wages between groups with different ex-ante skill levels (typically education). Out of these papers, the closest to our work is Caunedo et al. (2023) who allows for heterogeneity: some types of capital substitute for labor while others are complements. Overall, our findings provide direct evidence consistent with the view in Autor et al. (2003) regarding the potential for automation technologies to substitute for labor in routine cognitive tasks.

1 A model with automation and skill displacement

We motivate our empirical analysis with a model that captures the key channels through which technology affects earnings of individual workers. Technological progress is embodied in new vintages of capital. Our model builds on Acemoglu and Restrepo (2018) in that workers perform different tasks and Caunedo et al. (2023) in that capital improvements are specific to a particular set of tasks and can be either complements or substitutes. Thus, our model allows for both labor-saving as well as labor-augmenting technology improvements, depending on whether these improvements reflect capital-specific technological changes to routine or non-routine tasks performed by workers. In addition, we allow for the possibility that workers have skills that are specific to a specific technology vintage. As a result, some workers may get left behind as their skills and expertise are not fully transferable across vintages of labor-augmenting technologies, as in Violante (2002).

1.1 Setup

The model has two periods; in the second period there is a stochastic improvement in technology, which we model as a fall in the (quality-adjusted) price of capital. To economize on notation we omit time-subscripts and denote log growth rates across periods as Δ. Unless otherwise noted, discussion refers to pre-shock variables. Appendix A contains detailed derivations and discussion, including about potential sensitivity of our analysis to specific modeling assumptions.

Output Aggregate output is produced by a continuum of competitive industries,

$$\bar{Y} = \left(\int_k Y(k)^{\frac{\chi-1}{\chi}} dk \right)^{\frac{\chi}{\chi-1}}. \tag{1}$$

There is free entry of all firms in the production sector. In equilibrium, the price of each industry's output is equal to its marginal cost and firms make zero profits. To simplify exposition, we will be focusing on a given industry and drop the k subscripts unless needed.

The output of a single industry Y (the numeraire good) is a CES aggregate of a large number of intermediate tasks

$$Y = \left(\sum_{j=1}^{J} y(j)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}}. \tag{2}$$

The parameter $\psi > 0$ indexes the elasticity of substitution across tasks, as well as the absolute value of the demand elasticity for each task output. Task j is produced using capital $k(j)$ and labor $l(j)$

$$y(j) = \left((1-\gamma_j)\, k(j)^{\frac{\nu_j - 1}{\nu_j}} + \gamma_j\, l(j)^{\frac{\nu_j - 1}{\nu_j}} \right)^{\frac{\nu_j}{\nu_j - 1}}. \tag{3}$$

Our notion of capital incorporates not only machines used in production but also process methods or software. The parameter $\nu_j > 0$ determines the elasticities of substitution between $l(j)$ and $k(j)$ when performing each task and $\gamma_j \in (0,1)$ governs factor shares.

We allow for both labor-saving and labor-augmenting technology improvements. Automation (labor-saving) technologies are a substitute for routine tasks performed by workers; for these tasks, improvements in the quality of capital will lower its effective price and therefore lead to greater degrees of automation. Labor-augmenting technologies complement workers' non-routine tasks; for these tasks, improvements embodied in capital reflect improvements in the tools that workers use in their job, which in principle should enhance their productivity. To capture this distinction, we partition the set of tasks (2) into routine and non-routine tasks, $J = J_R \cup J_N$.

There are two distinctions between routine ($j \in J_R$) and non-routine ($j \in J_N$) tasks. First, the elasticity of substitution between capital and labor, ν_j, is lower for nonroutine tasks than for routine tasks. We assume that γ_j and ν_j is constant within the set of routine and non-routine tasks, with $\nu_N < \nu_R$. Second, due to skill displacement, new vintages of labor-augmenting capital can change the effective quantity of labor supplied to non-routine tasks.

Capital and Technology Firms chose the demand for capital deployed in each task taking prices $q(j)$ as given. The supply of capital suitable for task j is perfectly elastic. Technological innovation consists of declines in the (quality-adjusted) price of capital,

$$\Delta \log q(j) = -\varepsilon(j). \tag{4}$$

Technological progress can simultaneously affect multiple tasks: $\varepsilon \equiv [\varepsilon_1 \ldots \varepsilon_J]$ is a vector of weakly positive random variables jointly distributed according to $f(\varepsilon)$. Technology improvements are independent and identically distributed across industries.

Labor Supply In the initial period, there is a continuum of incumbent workers of measure I who supply labor across the different tasks j. The aggregate supply of labor in task j is given by:

$$L(j) = \int_0^I l(i,j)\,di, \tag{5}$$

where $l(i,j)$ is the number of efficiency units of labor supplied by incumbent worker i in task j, and $L(j)$ is the aggregate supply of labor in task j from workers who are incumbents in the first period. Each worker is associated with a single occupation $o(i)$. More precisely, workers in occupation o supply labor in only a small number of routine and nonroutine tasks: we denote by $J_o \subset J$ the set of tasks performed by occupation o. As in Acemoglu and Restrepo (2022), each task can only be performed by a single occupation, so $\bigcup J_o = J$ and $\bigcap J_o = \emptyset$. For incumbent workers, we assume that frictions to switching occupations are sufficiently large that her occupation is fixed in both periods. For simplicity, we assume that in the first period each worker has the same skill in each task she performs $l(i,j) = \bar{l}(i)$, $\forall j \in J_{o(i)}$–i.e., initial productivity differences across workers only reflect absolute advantage–and the distribution of $\bar{l}(i)$ is iid across occupations.

We make the following assumptions about changes in aggregate and individual labor supply between the two periods. First, we allow for skill displacement at the individual level: the adoption of a new vintage of technology leads to a change not only in wages but also in the number of efficiency units of an incumbent worker's human capital. Specifically, worker i's output in task j at time t satisfies

$$\Delta \log l(i,j) = \begin{cases} -\beta I[j \in J_N]\,\varepsilon(j) + u_{i,j} - \log l(i,j) & \text{w/ prob. } \omega\, I[j \in J_N]\,\varepsilon(j) \\ -\beta I[j \in J_N]\,\varepsilon(j) & \text{otherwise} \end{cases}, \tag{6}$$

where ω is a positive constant indicating the strength of the skill displacement effect, and $u_{i,j}$ is an i.i.d. draw from the distribution of $\log l(i,j)$ for incumbent workers in the same occupation in the initial period.

Our specification of vintage-specific human capital in (6) captures the intuition in Violante (2002), in which the fraction of prior skills that transfer to the new vintage is decreasing with the size of the technological improvement. As a result, larger shifts in the technology frontier are likely to generate greater amounts of skill displacement among incumbent workers (the first term). The second term in (6) is a force for redistribution: the most productive workers in the current vintage are most likely to experience declines in their productivity if technology changes. This term captures the idea that incumbent workers that are highly skilled in the prior vintage of technology will, on average, will be at a relative disadvantage when a new vintage is introduced.

Second, we allow for the aggregate supply of labor for task j to also vary on the extensive margin. Rather than explicitly modeling the flow of workers, we assume that aggregate labor supply satisfies

$$\Delta \log L(j) = \bar{\zeta} + \zeta_j \Delta \log w(j)\,, \tag{7}$$

Equations (7) and (6) together imply that new entrants are, on average, more highly skilled in the new technology than incumbents.

1.2 Worker earnings growth and technology exposure

The level of wage earnings for an individual worker is equal to the total compensation for the tasks she supplies. Thus, worker i's earnings growth evolves according to

$$\Delta \log W(i) \approx \left[\frac{\psi - \nu_R}{\nu_R + \zeta_R}\,\Gamma_R\right] \xi^R(i) + \left[\frac{\psi - \nu_N}{\nu_N + \zeta_N}\,\Gamma_N - \beta\right] \xi^N(i) - \omega\left[\log \bar{l}(i) - \int \log \bar{l}(i)\,dF(i)\right] \xi^N(i) \tag{8}$$

$$+\left[A_N + [A_R - A_N]\,\theta(i)\right]\Delta\log X,$$

where X is aggregate productivity (number of units produced per of dollar of input expenditure), $F(i)$ is the cumulative distribution function of $\log\bar{l}(i)$ across workers, and

$$\Gamma_j \quad\equiv\quad \frac{\partial\log p(j)}{\partial\log q(j)} = \frac{(\nu_j+\zeta_j)\kappa_j}{\psi+\zeta_j+(\nu_j+\zeta_j)\kappa_j} \tag{9}$$

captures the elasticity of marginal cost of each task $p(j)$ with respect to improvements in capital specific to that task $q(j)$, and $\kappa(j)$ is the ratio of expenditures on capital and labor in task j, and

$$A_j \quad\equiv\quad \frac{\partial\log w(j)}{\partial\log X} = \frac{(\chi-\psi)(1+\kappa_j)}{\psi+\zeta_j+(\nu_j+\zeta_j)\kappa_j} \tag{10}$$

equals the elasticity of task-level wages $w(j)$ to aggregate productivity X. We denote by $\theta(i)$ the share of labor compensation to worker i due to performing the routine tasks, the weight $s(i,j)$ denote the share of task j in worker's wage,

$$\theta(i) \equiv \sum_{j\in J_R} s(i,j), \qquad s(i,j) \equiv \frac{w(j)\,l(i,j)}{W(i)}\,I(j\in J_o) \tag{11}$$

and

$$\xi^R(i) \equiv \theta(i)\sum_{j\in J_R}\tilde{s}^R(i,j)\,\varepsilon(j) \quad\text{and}\quad \xi^N(i) \quad\equiv (1-\theta(i))\sum_{j\in J_N}\tilde{s}^N(i,j)\,\varepsilon(j), \tag{12}$$

denote the worker's exposure to labor-saving and labor-augmenting technologies, respectively; where $\tilde{s}^R$ and $\tilde{s}^N$ are the task weights normalized to sum to one for routine and nonroutine tasks. Last, we restrict attention to the case in which κ_j and ζ_j are the same across all routine and nonroutine tasks. Appendix A contains additional details on the derivation of (8).

Ideal Exposure Measure. Equation (12) states that a worker's exposure to labor-saving (automation) technologies is a function of the contribution of routine tasks to her compensation θ and the extent to which specific technological improvements are related to her routine tasks. Similarly, the worker's exposure to labor-augmenting innovation is related to the contribution of non-routine tasks to her compensation $1-\theta$ and the distance between technological improvements and her non-routine tasks. To fully implement (12) we would need data on compensation at the task level and the task-specific skills of individual workers. This data is not available, so we construct our exposure measures at the occupation, rather than at the individual worker, level.

The fact that workers in occupations that perform routine tasks are, on average, more exposed to labor-saving technologies is already emphasized by Autor et al. (2003). Our emphasis is distinct: the share of routine tasks she performs may measure the potential for her tasks to be automated, but the extent that these tasks are *actually* automated can be inferred by the distance between her tasks $\tilde{s}^R$ and $\tilde{s}^N$ and the degree of technological progress $\varepsilon(j)$.

Testable Predictions. Equation (8) forms the pillar of our empirical analysis. The first term in (8) captures the direct effect of automation on worker earnings—technology improvements $\varepsilon(j)$ in capital specific to routine tasks. The sign of the elasticity of task compensation $w(j)$ to quality improvements $\varepsilon(j)$ depends on the sign of $\psi - \nu_j$, where ν_j is the elasticity of substitution between capital and labor and ψ, which is equal to the elasticity of demand for that task (Hicks, 1932). If $\nu_j > \psi$, the high substitutability between cheaper capital and labor more than offsets the increase in labor demand for task j, which leads to lower wages for task$j$. By contrast, if $\nu_j < \psi$, demand for the output of task j is sufficiently elastic to lead to higher labor demand and therefore wages. For routine tasks $j \in J_R$, a reasonable parametrization is that $\nu_R >> \psi$, so improvements in automation technologies will lowers worker earnings. The second term in (8) captures the effect of labor-augmenting technologies on the compensation of labor assigned to the same task. Its sign is ambiguous, as it depends on the relative strength of two forces: first, whether the complementarity across tasks ψ is greater or smaller than the complementarity between capital and labor in non-routine tasks ν_N and second on the degree of skill displacement β among incumbent workers.

The third term in (8) captures the redistributive effects of skill displacement: the same improvement in labor-augmenting technology can in principle lead to winners and losers among incumbent workers. Recall that equation (6) implies that workers' current level of skill need not transfer perfectly to the new technology. As a result, for two workers in the same occupation, those workers with higher levels of initial skill (or income) are more likely to be displaced by the arrival of new technologies. Regardless of the sign of the average response, the model implies that workers with greater levels of skill in the previous technology are more likely to experience earnings declines following labor-augmenting technological change relative to their less skilled coworkers in the same occupation.

The last term in (8) captures the impact of technology spillovers on firm labor demand. Worker earnings grow with the change in productivity $\Delta \log X$, and worker exposures depend on their individual characteristics (skills and, most importantly, occupation). The strength of this channel depends on the elasticity of demand for industry output χ relative to the elasticity of demand for tasks ψ. When calibrating the model in Section 4, we impose that $A_R = A_N$. Therefore, this last term in (8) is fully absorbed by industry-specific time fixed effects. We revisit spillovers and the aggregate implications of our model in Section 4.

2 Measuring Workers' Technology Exposure

We next describe how we measure exposure to labor-saving and labor-augmenting technologies.

2.1 Methodology

Constructing an empirical analogue of equation (12) entails several challenges that we discuss next.

Classifying tasks based on their capital/labor complementarity

Most occupations perform a mix of tasks, and each task can be substituted or complemented by capital. To construct an accurate measure of exposure to labor-saving and labor-augmenting technologies, we therefore need to classify worker tasks as either substituted or complemented by capital. To do so, we rely on two complementary ideas in the literature. The first is the notion of routine tasks (Autor et al., 2003; Acemoglu and Autor, 2011); the key idea here is that routine tasks are the tasks that can be *potentially* performed by machines. By contrast, non-routine tasks cannot be performed by machines alone; instead machines increase workers' productivity when performing these tasks. Second, we rely on the notion of skill-biased technical change, in which there are a certain type of tasks that require low levels of related skill that can be performed by machines, while capital is a complement to labor when performing high-skill tasks (Krusell et al., 2000; Goldin and Katz, 2008). To partition the set of tasks performed by a set of occupation into these groups, we rely on recent advances in generative AI (GPT4) applied to the description of job tasks a given occupation performs from the Dictionary of Occupational Titles (DOT).

In brief, we query GPT4 on whether a given task can be characterized as routine or non-routine. Appendix B.2 discusses this procedure and the data in more detail. We validate the resulting classification at the occupation level based on a measure of routine task intensity (RTI) constructed from Acemoglu and Autor (2011) occupational task scores. To compare the two, we calculate the average share of tasks at the occupation level that are classified as routine by GPT4. Panel A of Figure A.1 plots RTI versus the average share of routine tasks at the occupation level. We see that the two are highly correlated (81%) and the relation is approximately linear.

As a robustness check, we also construct an alternative task partition based on the degree of required experience—a proxy for low- vs high-skill. We validate our alternative measure using ONET's job zone classification, which assigns occupations a score between 1 (lowest) and 5 (highest) depending on the extent of their required preparation, which includes education requirements, related experience, and on-the-job training. As we see in Panel B of Figure A.1, there is a strong correlation (85%) between the share of tasks classified as high-experience by GPT4 and ONET's job classification.

Breakthrough innovations

Next, we need to identify major improvements in technology. Taken literally, technology improvements in the model are embodied in capital—they correspond to declines in the quality-adjusted price of capital goods. But the definition of capital can be broad: it can relate not only to physical capital (for example, machines) but also intangibles (e.g. production methods or software). To identify these innovations we will rely on patent data and follow the methodology of Kelly et al. (2021), henceforth KPST. KPST identify breakthrough innovations as those that are both novel (whose descriptions are distinct from their predecessors) and impactful (they are similar to subsequent innovations). In particular, KPST first create a measure of importance for each patent that combines novelty and impact and then define a 'breakthrough' patent as one that falls in the top 10% of the distribution of importance. KPST show that these breakthrough technologies are associated with increases in measured productivity both at the aggregate as well as the industry level. KPST provide several indices; given that the administrative data sample has a shorter time dimension, we use the breakthrough definition of KPST that relies on 5-year forward similarity.

Measuring similarity between occupation tasks and patents

We next measure in constructing an empirical analogue of (12) is to measure the distance between a given technology and the tasks performed by workers of a specific occupation. We briefly describe our approach here; Appendix B.1 contains details.

We represent each document (either a patent or a part of an occupation task description) as a weighted average of the set of word vectors x_k for the terms contained in the document:

$$\mathcal{X}_i = \sum_{x_k \in A_i} w_{i,k}\, x_k. \tag{13}$$

The weights $w_{i,k}$ are based on the 'term-frequency-inverse-document-frequency' (TF-IDF), which overweighs word vectors for terms that occur relatively frequently within a given document and under-weighs terms that occur commonly across all documents. Next, we calculate the cosine similarity between a breakthrough patent b and the routine or non-routine component of the task description of occupation o,

$$\rho^j(b,o) = \frac{\mathcal{X}_b}{||\mathcal{X}_b||} \cdot \frac{\mathcal{X}_o^j}{||\mathcal{X}_o^j||}, \qquad j \in \{R, N\}. \tag{14}$$

We perform two subsequent adjustments to (14). First, we remove yearly fixed effects to account for language and structural differences in patent documents over time; patents have become much longer and use much more technical language over the sample period. Second, we impose sparsity to focus only on high degrees of similarity. Specifically, after removing the fixed effects we set all patent–occupation pairs to zero that are below the 80th percentile in this fixed-effect adjusted similarity. Thus, only similarity scores sufficiently high in the distribution receive any weight. Last, we scale the remaining non-zero pairs such that a patent/occupation pair at the 80th percentile of yearly adjusted similarities has a score equal to zero and the maximum adjusted score equals one. We denote by $\tilde{\rho}^j(b,o)$ the adjusted similarity metric. We also repeat the exercise for the alternative classification of tasks into high- and low-experience.

In sum, we propose a method that uses a combination of word embeddings and TF-IDF weights in constructing a distance metric between a patent document (which includes the abstract, claims, and the detailed description of the patented invention) and the routine and non-routine component of detailed description of the tasks performed by occupations.[4]

2.2 Specific examples and validation

We next verify that our measure indeed captures exposure of workers to related technologies. First, we examine specific examples of breakthrough technologies and identify the most related occupations. Second, we validate our technology exposure measures using a large language model (GPT4).

Examples

A key advantage of our measure is that it allows us to isolate technologies that are likely to complement or substitute for workers in a specific occupation. As an example of labor-saving technologies, consider US Patent 5,309,803 for a "System including method and apparatus for cutting each layer of a double-layered roll of sheet to different lengths and widths" is the patent most closely related to the routine component of "Cutting and Slicing Machine Setters, Operators, and Tenders" and "Textile Cutting Machine Setters, Operators, and Tenders", two occupations that perform primarily routine tasks. Similarly, US Patent 5,518,574 for a "Form folding and gluing machine" is most closely related to "Adhesive Bonding Machine Operators and Tenders", another occupation that mainly performs routine tasks; while Patent 6,044,352 for " Method and system for processing and recording the transactions in a medical savings fund account" is most closely related to the routine part of the job description for "Insurance Claims and Policy Processing Clerks".

As an example of a labor-augmenting technology, consider US Patent 5,189,606 titled "Totally integrated construction cost estimating, analysis, and reporting system", which is the patent most closely related to the non-routine component of "Architects, Except Landscape and Naval", an occupation that performs only non-routine tasks. Another example is "Radio Operators", an occupation that performs mainly, but not exclusively, non-routine tasks. Among the patents most closely related to their non-routine tasks description are US Patent 5,123,112 for "Air-to-ground communication system" and Patent 5,212,804 for a " Communication system having multiple base stations and multiple mobile units". Similarly, Patent 5,687,093 for " Integrated system for

[4] Our methodology is conceptually related, though distinct, to the method proposed by Webb (2020), who also analyzes the similarity between patents and job tasks. Webb (2020) focuses on similarity in verb-object pairs in the title and the abstract of patents with verb-object pairs in the job task descriptions and restricts attention to patents identified as being related to robots, AI, or software. We use instead the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention. Besides the difference in scope, there are also differences in methodology: Webb (2020) uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings; we infer document similarity by using geometric representations of word meanings (GloVe) that have been estimated directly from word co-occurrence counts. Autor et al. (2022) build on the approach developed in this paper to isolate the role of technologies in creating demand for new tasks. Seegmiller, Papanikolaou, and Schmidt (2023) considers applications of the approach we develop here to economic history.

gathering, processing, and reporting data relating to site contamination" is most closely related to the non-routine component of the task description of "Environmental Engineers".

The examples above were focusing on occupations that performed mostly routine or non-routine tasks. However, given that our measure is constructed using a partition of a worker's tasks, it allows for some workers to be only partially exposed to some technologies. For instance, consider the occupation titled "Computer Programmers", which performs an even mix of routine and non-routine tasks. One of the patents most closely related to the routine parts of that occupation is US Patent 6,948,168, titled "Licensed application installer", which consists of a method for automatically installing software. By contrast, US Patent 5,806,053, titled "Method for Training a Neural Network" is closely related to the occupation's non-routine tasks. Another example is US Patents 5,797,126 for "Automatic theater ticket concierge" and 5,502,806 for a "Waiting line management system"; these two patents are most closely related to the routine and non-routine parts of the task description for "Ushers, Lobby Attendants, and Ticket Takers".

It is important to emphasize that each individual patent likely has a small impact on worker earnings; given that we are aggregating across thousands of patents, the time-series behavior of our measure will be driven by a collection of related technologies. For instance, one labor-displacive technology picked up by our measure is the rise of e-commerce—and more specifically the automatic fulfillment of retail purchase orders. Our measure indicates the period of the 1990s as featuring a significant surge in innovation related to the routine component of the tasks performed by order-fulfillment clerks. Examples of such breakthrough innovations early on include U.S. Patents 5,696,906 and 5,884,284 for "Telecommunication user account management system and method"; or Patent 5,627,973 for "Method and apparatus for facilitating evaluation of business opportunities for supplying goods and/or services to potential customers". Seegmiller et al. (2023) includes additional examples and a validation of our methodology.

Validation

Our operating assumptions thus far have been that technology that relates to routine tasks is likely a substitute for workers (labor-saving) while technology that relates to workers' non-routine tasks is likely to be a complement (labor-augmenting). To validate these assumptions we use GPT4: we focus on a random sub-sample of 10,000 breakthrough patents and the set of five occupations that are closest in terms of routine or non-routine tasks. For the set of occupations that are closest in terms of routine tasks, we query GPT4 on whether the technology in question is likely to substitute for the tasks performed by these occupations. For the occupations that are closest in terms of non-routine tasks, we query GPT4 on whether this technology is likely to increase the productivity of workers in these occupations. As a placebo, we repeat the same two queries for the five occupations that are least related in terms of their routine or non-routine tasks. Appendix B.4 contains further details.

GPT4 largely agrees with our classification (Appendix Table A.1). Specifically, for the five most closely related occupations in terms of their routine tasks, GPT4 agrees that this technology is likely to be labor-saving in 86% of the cases, while it only agrees with 4% of the cases for the least exposed occupations. Similarly, when focusing on the distance to non-routine tasks, GPT4 agrees in 83% of the cases that this technology is likely to be complementary to workers in the five most closely related occupations, while only in 10% of the cases for the least exposed occupations. Columns (2) and (4) report the results of the same validation exercise for our high/low required experience measure, which are largely comparable.

2.3 Worker technology exposure

Having validated our measures of technology exposures to labor-saving and labor-augmenting technologies, the final step is to create worker-level exposure metrics. Our data has detailed information on both the industry of a particular worker as well as the industry of origination of each patent. This allows us to exploit additional sources of variation: we can compare workers not only across occupations in the same industry, but also workers in the same occupation across industries. Letting b index breakthrough patents; $\mathcal{B}_{k,t}$ denote the set of breakthrough patents issued in industry k in year t; o denote occupations; and k denote industry at the NAICS4 level.

We construct the empirical analogue of equation (12) in the model. In particular, we create an index of exposure of workers in industry k and occupation o to technology at time t as

$$\xi^j(k,o,t) = \theta^j(o)\,\log\left(1 + \sum_{b\in\mathcal{B}_{k,t}} \tilde{\rho}^j(o,b)\right), \qquad j \in \{R, N\} \tag{15}$$

Our measure $\xi^j(k,o,t)$ varies over time t and industry k due to the differential arrival of breakthrough technologies across industries and it varies across occupations o as these breakthrough technologies have different levels of similarity with the tasks performed by each occupation. Specifically, $\xi(k,o,t)$ aggregates our patent-occupation similarity scores $\tilde{\rho}^j(o,b)$ across all breakthrough patents $\mathcal{B}_{k,t}$ issued to a firm in industry k in period t. We match each patent to the industry of the patent assignee at the 4-digit NAICS level using restricted access Census data. Given that just under half of the industry–occupation pairs have zero breakthrough patents in a given year, we apply a log transform to smooth out the skewed distribution. Following equation (12), we also multiply our measure with the shares θ^R_o and θ^N_o of routine or non-routine tasks performed by occupation o, respectively.

We next examine how our exposure measures vary across occupations. Figure 1 shows the distribution of our exposure measures by broad occupation categories. We see that workers employed in production and construction, are in general more exposed to automation than the average worker, along with some occupations in the transportation industry. Out of all the white collar jobs, office workers are most exposed to automation. By contrast, workers in high-skill service jobs, and especially business and STEM occupations are relatively more exposed to complementary technologies rather than automation. Figure 2 shows how these technology exposures vary across occupation wage levels. We see that the most exposed occupations to labor-saving technologies tend to be found in the middle of the income distribution, consistent with the prevailing view regarding job polarization in the United States (Autor and Dorn, 2013; Bárány and Siegel, 2018). By contrast, the exposure of occupations to labor-augmenting technologies is increasing in their average wages.

3 Technology Exposure and Worker Outcomes

Armed with a measure of workers' technology exposure, we next move to the main goal of the paper: understanding how exposure to major technologies shapes outcomes of individual workers.

3.1 Administrative employer-employee data

We combine employer-employee matched Social Security Administration (SSA) administrative data (the Detailed Earnings Records file) with survey information for a random sample of individual workers tracked by the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC). The SSA data include information on income and employer identification numbers (EINs) from Form W-2. The ASEC includes information on occupation as well as demographic information such as age and gender. We complement the data with information on industry and average wages from the Census Longitudinal Business Database (LBD) using the EINs. We limit the sample to individuals who are older than 25 and younger than 55 years old and to periods where the CPS interview date is within the past 3 years so that the occupation information is relatively recent.

Our sample contains approximately 2.8 million person-year observations spanning the period from 1981 to 2016. Approximately 54% of the sample is male and 34% of the observations correspond to workers with a four-year college degree. The median worker in the sample is 41 years old and earns approximately $50k per year in terms of 2015 dollars. The distribution of earnings is rather skewed: the average is equal to $66k while the 5th and 95th percentiles are equal to $16k and $152k, respectively. Importantly, much of these differences in pay exist within industry–occupation groups; on average, about 58 percent of the cross-sectional dispersion exists within industry–occupation cells. The dispersion in wage earnings within an occupation and industry has been growing steadily over time, in line with the overall increase in income inequality across all workers (Appendix Figure A.2). Appendix Table A.2 summarizes the sample and Appendix B.7 contains further details on sample construction.

In our empirical analysis, we will interpret this sizable heterogeneity in worker earnings within an industry–occupation cell as reflecting within-job skill differences across workers, consistent with the view in Goldin and Katz (2008). These within industry–occupation differences in wage earnings are only partly driven by firm heterogeneity: the firms that employ the workers at the top 5 percent of the within industry–occupation earnings distribution are approximately 0.45 log points more productive on average than the firms that employ the workers at the bottom 25 percent. However, the bottom panel of Appendix Figure A.2 shows that these between firm differences account for only a small fraction (about 16 percent) of the within industry–occupation earnings dispersion.

3.2 Empirical methodology

We estimate the following specification,

$$\Delta w^i_{t+h} = \gamma\, \xi^R_{i,t} + \delta\, \xi^N_{i,t} + c\, \mathbf{Z}_{i,t} + u_{i,t}. \tag{16}$$

The dependent variable in (16) is the growth in worker i's average W2 earnings over the next $h = 3, 5$ and 10 years, relative to the prior three years. To smooth out the impact of transitory earnings spikes, we follow Autor, Dorn, Hanson, and Song (2014) and Guvenen, Ozkan, and Song (2014) and consider the growth in average wage earnings, adjusted for life-cycle effects,

$$\Delta w^i_{t+h} \equiv w^i_{t+1,t+h} - w^i_{t-2,t} \qquad \text{and} \qquad w^i_{t,t+h} \equiv \log\left(\frac{\sum_{j=0}^{h} \text{W2 earnings}_{i,t+j}}{\sum_{j=0}^{k} D(\text{age}_{i,t+j})}\right). \tag{17}$$

Our two independent variables of interest are the technology exposure measures $\xi^j_{i,t}$ that we can assign to an individual worker i at time t based on her occupation $o(i)$ and the (NAICS 4-digit) industry $k(i,t)$ of the worker in each year (15). The time t is determined by when the patent is issued to a firm in industry k. The actual implementation of the technology can occur earlier, if the firm developing the technology deploys it before the patent is issued, or later, if there are adoption lags. Given that we lack data on the actual implementation of these technologies, we use the issue year as our baseline, and vary the horizon h over which we examine the relation to worker earnings.

Equation (16) estimates a version of model equation (8). The coefficients γ and δ are identified by comparing future earnings for a highly exposed worker to less exposed workers that are the same age, have the same level of earnings and prior earnings history, and either work in the same industry in a different occupation, or in the same occupation in other industries. Specifically, the vector of controls **Z** includes flexible non-parametric controls for worker age and the level of past worker earnings as well as recent earnings growth rates. In addition, we include different combinations of year, occupation, worker prior income, and industry fixed effects, depending on the specification.[5]

Equation (17) focuses on the first three terms of model equation (8). Our baseline specification interacts occupation, worker prior income, and industry fixed effects with calendar year to account for occupation- or industry-specific time trends. This saturated specification allows us to partial out

[5]We construct controls for worker age and lagged earnings by linearly interpolating between 3rd degree Chebyshev polynomials in workers' lagged income quantiles within an industry-age bin at 10-year age intervals. In addition, to soak up some potential variation related to potential mean-reversion in earnings (which could be the case following large transitory shocks), we also include 3rd degree Chebyshev polynomials in workers' lagged income growth rate percentiles, and we allow these coefficients to differ by gender as well as past income levels based on five gender-specific earnings bins.

sources of time-series variation that affect specific groups of workers in a given occupation, industry, or income level. The cost of saturating our specification is that these fixed effects largely absorb the last term in equation (8) which captures productivity spillovers across workers.

3.3 Average efffects

We begin by estimating a version of equation (16) in which the estimated coefficients γ and δ do not vary across workers. This restricted version maps into equation (8), but with the third term averaging to zero across workers. Table 1 plots the estimated coefficients γ and δ for horizons of $h = 3, 5$ and 10 years for different combinations of fixed effects. To conserve space, we focus our attention on the last column, which reports the estimates corresponding to the most saturated specification that includes occupation and industry fixed effects both interacted with calendar year.

Our technology exposure measures ξ^R and ξ^N are both significantly negatively related to workers' subsequent earnings growth. The magnitudes are not small: an increase in the workers' exposure to automation ξ^R from the median to the 90th percentile is associated with approximately a 2.2 to 2.9 percentage point decline in worker cumulative earnings over a period of three to ten years. Exposure to complementary technologies ξ^N is also associated with earnings declines for workers, though the magnitudes are still sizable but somewhat smaller—ranging from 1 to 1.7 percentage points, depending on the horizon. In both cases, magnitudes increase with horizon suggesting these are persistent effects. Appendix Table A.3 shows the coefficient estimates that correspond to our alternative measure of technology exposure using the high/low required experience classification created in the second half of section 2.1. We see that these estimates are very comparable to our baseline measure.

Here, there are several points worth discussing. First, the estimated coefficient γ is negative, which is consistent with the idea that technologies that are related to routine tasks act as a substitute for workers, that is $\nu_R - \psi > 0$ in the model. Second, the estimated coefficient δ is also negative, though significantly smaller in magnitude than γ. In the model, the sign and magnitude of this coefficient depends on the complementarity between capital and labor in non-routine tasks, $\nu_N - \psi < 0$ and the level of skill displacement among incumbent workers β. Third, the implied magnitudes are sizeable and comparable to the observed variation in average log earnings growth over the business cycle. Specifically, in US administrative data for the 1998-2019 period, the standard deviations of 1 and 5 year changes in average log earnings growth are around 1.5 and 2.9 pp, respectively, and the peak-to-trough change in average 5 year earnings growth in the Great Recession is 5.2 pp.[6]

Last, we emphasize that this empirical specification mainly reveals workers' direct technology exposure; it does not include the effect of spillovers from productivity improvements. Our preferred

[6] These calculations are based on annual data from the Global Repository of Income Dynamics.

specification in column (4) has industry-year fixed effects, which largely absorb the last term in equation (8) capturing spillovers from technologies to workers performing unaffected tasks. Comparing columns (1) to (3) of Table 1 to our preferred specification in column (4), we see that including industry times calendar year fixed effects has a modest negative impact on the estimated γ coefficient, whereas the estimate of δ becomes significantly more negative—almost double the magnitude in some cases. This finding is consistent with our conjecture above that the industry-year fixed effects in our specification likely absorb the potential positive benefits of technology exposure on worker wages through increases in industry productivity. We revisit this issue in Section 4. By contrast, allowing for occupation-specific trends in earnings growth (occupation interacted with year fixed effects) has only a modest quantitative impact on our results, suggesting that this is not the dominant source of variation that identifies our estimates.

3.4 Heterogeneity across sectors and job types

We next allow the estimated coefficients γ and δ to vary across different sectors or job types. Panel A of Figure 3 compares the coefficients across workers in manufacturing (broadly defined by 2-digit NAICS codes 11-33) versus the remaining industries (2-digit NAICS codes 42-92). Panel B compares across occupation types, depending on whether they primarily emphasize manual, cognitive or interpersonal skills.

We see that the estimated coefficient γ capturing the impact of labor-saving technologies on earnings is comparable both across manufacturing and service industries, as well as between manual and cognitive-intensive jobs. By contrast, workers in occupations emphasizing interpersonal tasks are essentially unaffected by our labor-saving technology exposure measure. This pattern suggests that the displacement of workers in response to labor-displacive technologies is not purely a blue-collar worker phenomenon; rather it is increasingly present in white-collar occupations, supporting with the view in Autor et al. (2003) regarding the potential for automation technologies to substitute for labor in routine cognitive tasks.

By contrast, the estimated coefficient δ capturing the impact of labor-augmenting technologies is only negative, and statistically different from zero for workers employed in the non-manufacturing sector or in cognitive-intensive jobs. Interpreted through the lens of this model, this fact would suggest a higher level of vintage capital specificity in white collar than blue collar jobs.

3.5 Heterogeneity by worker skill

We now allow the estimated coefficients γ and δ in equation (16) to vary with observable measures of worker skill. Recalling equation (8), we see that the first term, which corresponds to the estimated coefficient γ, should be constant across workers. By contrast, the second and third term together imply that the estimated coefficient δ should become increasingly more negative for higher levels of initial skills $\bar{l}(i)$. Though worker skill is not directly observable in our data, we can rely on observable measures: a worker's age, her level of wage earnings relative to their peers, and her education level.

Age

We first allow both coefficients γ and δ to vary with workers' age. We use worker age as a proxy for the workers' level of skill in the tasks performed by their occupation, rather than job tenure, since we do not observe the full occupation history of the worker. Panel A of Table 2 reports the estimated coefficients γ and δ across workers in different age groups—25 to 35, 35 to 45, and 45 to 55. To conserve space, the we report results for horizons of five years and scales the coefficients so that they correspond to a shift in exposures from the median to the 90th percentile.

We see that the coefficient estimates are in line with our model's prediction. Specifically, the estimated coefficient γ capturing the impact of labor-saving technologies is consistently negative across workers in all age groups. The point estimates of γ range from 1.8 to 3.2 percentage points, but there is no discernible pattern across age groups. By contrast, the estimated coefficient δ shows a strong age pattern: the estimate is essentially zero for the youngest workers but is negative and statistically significant for older workers (workers in either the 35–45 or in the 45–55 group). The difference in the estimated coefficients δ between the group of oldest and youngest worker is approximately equal to 1.7 percentage points, and is both statistically and economically significant. Appendix Figure A.5 shows that this pattern persists across horizons, while Panel A of Table A.4 shows that we obtain quantitatively similar estimates of γ and δ across groups using our alternative technology exposure measures that rely on classifying tasks based on prior experience.

Relative earnings

We next allow the coefficients γ and δ to vary with worker's earnings over the last year relative to her peers—workers in the same occupation and industry. To the extent that more highly skilled workers are paid more relative to their peers, our model implies that we should see the estimated coefficients δ decline with a worker's prior income. By contrast, our model has no implications about the extent to which γ varies with prior income levels.

Panel B of Table 2 plots the resulting estimates of γ and δ. Examining the figure, we see that the estimated coefficient γ corresponding to the impact of labor-saving technologies on worker earnings is consistently negative for all income groups, but does not vary significantly across income levels. By contrast, the coefficient δ corresponding to labor-augmenting technologies is only significantly negative for the most highly paid workers—workers in the top half of the within occupation-industry pay distribution, and especially the top quartile. In particular, an increase in our labor-augmenting exposure measure from the median to the 90th percentile is associated with a 1.1 percentage point decline in earnings for workers in the 50th to 75th percentile group, a 2.1 percentage point decline in earnings for workers between the 75th and the 95th percentiles, and a 4.7 percentage point decline for workers in the top 5 percent of the relative pay distribution. By contrast, the estimated coefficients are still negative but much smaller in magnitude and not statistically significantly different from zero for workers below the median.

Education

Last, we allow the coefficients γ and δ to vary with the worker's education. Existing work has emphasized that much of technological change is skill biased, where skill is often defined as worker education (see, e.g. Goldin and Katz, 2008). However, to the extent that worker education correlates with her level of her skill operating the current vintage of capital, but does not relate to the degree to which her skills are transferable across technology vintages, our model would imply that the estimated coefficient δ is more likely to be negative for college-educated workers than for workers without a college degree. To this end, we next compare whether the earnings growth of workers with and without a four-year college degree respond differentially to the same increase in their occupation-industry technology exposure. This comparison exploits both between- as well as within-cell variation in the college share.

Panel C of Table 2 shows the estimated coefficients γ and δ across these two groups of workers. As before, the estimated coefficient δ is larger for the more highly-skilled workers: an increase from the median to the 90th percentile in our exposure to labor-augmenting technologies is associated with 1.6 vs 1.0 percentage point decline for workers with and without a college degree, respectively. However, the difference is much smaller in magnitude, even though it is statistically different from zero. As before, we find no meaningful difference in our estimate of γ across these two groups. Panel C of Appendix Table A.4 illustrates that our estimates are comparable when using our alternative exposure measure based on the high/low skill classification of individual tasks.

Robustness

Our key finding that the earnings of top earners are more sensitive to improvements in labor-augmenting technologies are robust to several changes in methodology. First, Panel B of Appendix Table A.4 shows that the results are quantitatively very similar when we use our alternative measure of exposure to labor-saving and labor-augmenting technologies that rely on classifying tasks based on prior experience. Second, Appendix Table A.5 shows that our baseline results (reproduced in column one) are robust to alternative methods of ranking workers (columns two to seven). In particular, one concern is that ranking workers based on their prior earnings may be mechanically identifying workers that are recently hired, since those workers will have received wages for only part of the year. Thus, in column (2) we report estimates when we drop workers hired within the last year, while in column (3) we rank workers based on their average income over the last two years relative to their peers. The point estimates are very comparable to our baseline specification, suggesting that this is not a major concern.

Another concern is that workers' income is a very noisy measure of their level of skill in the current technology. To address this concern, we next remove some of the heterogeneity in worker earnings that may be unrelated to worker skill. Specifically, in column (4) we compute worker wages relative to the average wage offered by their employer; in columns (5) we rank workers based on their residual wage earnings net of occupation, industry, commuting zones, age, and gender fixed effects, all interacted with calendar year. Removing commuting zone (interacted with year) effects allows us to remove the component of pay that may be related to local labor scarcity, or monopsony power on the part of firms and workers, while removing age bin by gender fixed effects accounts for the fact that older workers tend to be more highly paid. Column (6) combines both adjustments from columns (4) and (5). Last, in column (7) we also compute wage earnings residuals net of the worker's unionization status interacted by calendar year. Doing so reduces the sample dramatically, since only approximately one in five workers in our sample provide an answer to the unionization question on the CPS. These adjustments produce estimates comparable to our baseline estimates.

3.6 Technology exposure and job loss

Our results so far have focused on the average response of worker earnings to our technology exposure measures. However, these average responses may mask considerable heterogeneity in ex-post outcomes for workers. Part of this heterogeneity is predictable using ex-ante characteristics: we saw that the wage earnings declines in response to labor augmenting technologies are concentrated on older and more highly-paid workers. However, there may also be heterogeneity in ex-post outcomes even conditioning on similar ex-ante characteristics. For instance, our estimates cannot distinguish between each worker experiencing a 2 percentage point decline in earnings or 10 percent of the workforce experiencing a 20 percentage point decline due to job loss.

We next explore the extent to which the wage earnings declines we document in response to our technology exposure measures are ex-post concentrated on a subset of workers. Specifically, we shift attention to large losses in worker earnings as our outcome of interest: we re-estimate (16) but now replace the dependent variable with a dummy capturing involuntary exit, which takes the value of one if the worker has switched employers at any point between time t and $t + h$ and at the same time experienced growth in worker earnings below the 20th percentile of earnings growth, calculated among all workers in that time period. The unconditional probability of involuntary exit in our sample is 17 percent, which implies that 85 percent of the workers with earnings growth in the bottom 20th percentile also separated from their employer.

Figure 4 plots the estimated coefficients γ and δ. In panel A we allow these coefficients to vary with the worker's age; in panel B they vary with the worker's prior income relative to her peers as in the previous section. Examining the top panel (A) of the figure and focusing on the red bars, we see that our labor-saving technology exposure measure is both statistically and economically significantly related to measures of worker (involuntary) exit. The magnitudes of the estimated coefficients are not small: an increase in ξ^R from the median to the 90th percentile is associated with a 0.9 to 1.7 percentage point increase in the likelihood of the worker experiencing involuntary exit—which is economically significant given the unconditional probability of 17 percent. As before, we do not see a strong pattern across age groups. By contrast, when we examine the response to our labor-augmenting measure of technology exposure (the blue bars) we see that, the magnitudes are lower on average by approximately 50 percent, but they are significantly more concentrated on the older workers: for a worker aged 45 to 55 years, an increase in labor-augmenting technology exposure from the median to the 90th percentile is associated with a 1 percentage point increase in the likelihood of involuntary exit—compared to essentially zero for workers aged 25 to 35.

The bottom panel (B) of Figure 4 shows a similar pattern when we condition on the worker's prior income relative to her peers. Focusing on the red bars showing the estimated coefficient γ, we see that an increase in exposure to labor-saving technologies ξ^R from the median to the 90th percentile is associated with a 1 to 2 percentage point increase in involuntary loss for workers; the magnitudes are weakly decreasing with workers' prior income. By contrast, the blue bars indicate that our labor-augmenting technology exposure measure is associated with a much more concentrated increase in the likelihood of exit. In particular, an increase in ξ^N from the median to the 90th percentile is associated with 1.7 and 3.8 percentage point increase in the likelihood of exit for workers in the 75th to 95th and over the 95th percentile of prior earnings. By contrast, the estimated coefficients are essentially zero for workers below the 75th percentile in terms of prior earnings.

In sum, we see that our measures of direct technology exposure are significantly associated with measures of (likely involuntary) worker job loss. The magnitudes are significantly larger for our labor-saving exposure measure, which is consistent with the idea that automation is associated with job loss. Given our definition of job loss, the expected growth rate of cumulative earnings for these workers is -94.1 log points (Table A.2). Thus, our estimates imply that the increased likelihood of job loss can account for between one half to more than two-thirds of the response of mean worker earnings to labor-saving technologies in Table 2. Perhaps more surprisingly, however, we find that our labor-augmenting exposure measure also predicts job loss for the most skilled workers, a finding which is consistent with the idea of skill displacement in the model, even though the model does not actually have an extensive margin of employment. In fact, for the workers at the top of the earnings distribution, the increased likelihood of job loss can account for almost three-quarters of the overall decline in mean earnings. Appendix Figure A.6 illustrates that using our alternative measure of labor-saving and labor-augmenting technologies—based on the characterization of tasks into high- and low-skill—leads to quantitatively very similar results.

3.7 Discussion

The results in the previous section reveal a significant degree of heterogeneity in the response of worker earnings to our measure of labor-augmenting technologies. Consistent with our model, this heterogeneity is related to observable measures of worker skill in the existing technologies—our estimated coefficient δ in in our empirical specification (16) is larger in magnitude for older, more highly educated, and more highly paid workers relative to their peers. Importantly, this heterogeneity in the estimated coefficient δ is significantly larger than the average effect: the response of earnings for top workers to the same improvement in labor-augmenting technologies is almost four times larger than the magnitude of the average coefficient. By contrast, we find no evidence for significant heterogeneity in how earnings respond to labor-saving technologies—the coefficient γ in (16) are essentially flat across age or prior income. As we relate our findings to the existing literature, there are several points worth discussing.

Capital-skill complementarity The higher response of more highly-paid or college-educated workers to measures of labor-augmenting technologies may be puzzling viewed through the lens of capital-skill complementarity. Under this view, worker skill translates to general skills that are transferable across technologies. If skill is an immutable characteristic of the worker, and skill is complementary to technology, we would expect to see that more highly-skilled workers experience either lower wage declines or wage increases in response to labor-augmenting technologies. We instead emphasize an alternative view, that worker skill is specific to a particular technology vintage. Are our results inconsistent with capital-skill complementarity? Not necessarily: the existing evidence in favor of capital-skill complementarity relates to skill differences across occupations; these between-occupation differences in skill likely reflect differences in general skills. By contrast, we primarily focus on within occupation and industry measures of skill, which likely reflect primarily variation in specific skills.

Directed Innovation. One potential concern is that part of the patterns we are capturing (our estimates of γ) reflect the endogenous development of labor-saving technologies targeted towards workers (occupation–industry cells) that have become more 'expensive' relative to their marginal productivity. If that is the case, we would expect that past wage growth would positively predict each one of our technology exposure measures. To explore if this is indeed the case, we compute the partial correlation between ξ_{t+5+h} and wage growth Δw^i_{t+5}, where both variables are orthogonalized with respect to the right-hand side variables of equation (16) using our most saturated specification and the exposure to labor-augmenting technologies (in the labor-saving case, and vice versa for the labor-augmenting case). To ensure that there is no overlap between the time the technology is implemented and the period over which we measure wage growth, we examine horizons h of three to seven years—over 80 percent of the patents in our sample have a 2-year lag between the time the application is filed to when the patent is approved, and the median lag is 3 years.

Examining Figure A.4 and focusing on the top panel corresponding to leads of the labor-saving technology exposure ξ^R, we see that there is little evidence that wage growth positively predicts our labor-saving technology exposure measure. The bottom panel of Figure A.4 shows that past wages are essentially unrelated to future realizations of our labor-augmenting technology exposure measure ξ^N. This lack of predictability of future technology exposure by past wage growth helps alleviate a related concern that the development of labor-augmenting technologies is targeted to worker groups that have experienced wage increases due to increased human capital accumulation. We obtain very similar patterns when we use our high/low skill definitions of labor saving and labor augmenting technologies. We conclude that even though innovation effort could in principle be directed to certain worker groups, the resulting innovation outcomes are sufficiently random to be not predictable by past worker wages.

Endogenous adoption of new technologies. A related concern is that the adoption of automation technologies is endogenous within an industry–occupation cell; if firms can target technology adoption to specific workers and the cost of adoption does not vary across workers, we would expect that they would target workers whose wages are high relative to their marginal productivity. If the dispersion on wages within industry–occupation cells is imperfectly related to worker productivity, then we would expect to find that the estimated coefficient γ would be larger for more highly paid (and perhaps older) workers. By contrast, we find that our estimated coefficients γ are economically comparable across workers with different age or prior income.

That said, we should emphasize that our results should not be used to reject a model of endogenous adoption for several reasons. First, it is possible that not only the benefit, but also the cost of adopting a technology is larger for workers that are paid more relative to their marginal product than other workers; for instance, these workers may be more resistant to the adoption of new labor-saving technologies. Second, it is possible that variation in workers' bargaining power is not the key driver behind the (sizeable) dispersion in wage earnings within industry–occupation cells. If anything, the correlation may be negative: we find that workers at the top of the relative earnings distribution within industry–occupation cells are more likely to be employed at more productive firms: firms employing the highest-paid workers in an industry–occupation cell have approximately 45% higher labor productivity (0.45 log point units) than firms employing the workers at the bottom of the income distribution. Since productive firms appear to exert market power in the labor market (Gouin-Bonenfant, 2020; Seegmiller, 2021), workers in these firms may be less likely to have higher bargaining power relative to workers in less productive firms.

4 Aggregate Effects of Technology Exposure

So far, we have been focusing on estimating the direct effect of technology exposure on worker earnings. However, this direct effect is only a partial description of how technology affects workers in the model for two reasons. First, our worker level regressions reveal the impact of technology on the earnings of incumbent workers, which is not necessarily equal to the impact of technology on all workers in a given occupation if only the incumbent workers suffer from skill displacement. Section 4.1 focuses on the distinction between outcomes for incumbent workers and aggregate outcomes at the occupation by industry level. Second, wage earnings of workers unaffected by technology improvements increase due to increases in labor demand as productivity improves. In Sections 4.2 and 4.3, we quantify these spillovers. We estimate the model using GMM in Section 4.4.

4.1 Employment and wages

Here, we explore the distinction between outcomes for incumbent workers compared to outcomes for all affected workers as a group, which includes new entrants. The distinction between incumbent workers and new entrants is relevant because of our assumption of vintage-specific human capital in operating labor-augmenting technologies: incumbent workers are displaced whereas new entrants are not.

Total employment at the occupation level o evolves according,

$$\begin{aligned}\Delta \log L(o) \approx \zeta_R \left[\frac{\psi - \nu_R}{\nu_R + \zeta_R}\, \Gamma_R\right] \xi^R(o) + \zeta_N \left[\frac{\psi - \nu_N}{\nu_N + \zeta_N}\, \Gamma_N\right] \xi^N(o) \\ + \Big(\zeta_R\, A_R \theta(o) + \zeta_N\, A_N (1 - \theta(o))\Big)\, \Delta \log X,\end{aligned} \tag{18}$$

The first two terms in equation (18) capture the direct effects of labor-saving and labor-augmenting technologies on employment in affected occupations. The sign of these terms depend on the relation between (ν_R, ν_N), and ψ. As long as $\nu_R > \psi$, employment should decline in occupations exposed to labor-saving technologies, whereas it should increase in occupations affected by labor-augmenting technologies if $\nu_N < \psi$. The response of average wages across all workers in the same occupation is given by (8) in the special case where $\beta = \omega = 0$. The response of the wage bill across all workers in the same occupation is equal to the change in wages plus the change in employment (18).

We estimate the empirical analogue of (18) using data from the Decennial Census and American Community Survey (see Appendix B.7 for details)

$$\frac{1}{h}\Big(\log(X_{o,k,t+h}) - \log(X_{o,k,t})\Big) = \gamma\, \xi^R_{o,k,t} + \delta\, \xi^N_{o,k,t} + c\, \mathbf{Z}_{o,k,t} + u_{o,k,t}. \tag{19}$$

The dependent variable in (19) is the annualized growth rate in employment, wages, or the total wage bill, aggregated at the occupation–industry–year level. We focus on a ten-year horizon $h = 10$ given the structure of the Decennial Census. Our vector of controls $\mathbf{Z}$ includes occupation and industry (NAICS4) fixed effects interacted with calendar year. Just like our worker-level regression in equation (16), these fixed effects absorb sources of time-series variation that affect specific groups of workers in a given occupation or industry; as before, they also absorb the last term in equation (18) capturing cross-occupation spillovers, which allows us to focus our attention to the first two terms. In addition we include controls for the share of female workers in the occupation–industry cell, and the logarithm of the average years of education, worker age, average wage, and employment at the start of the period.

Table 3 reports the estimated coefficients γ and δ. Focusing on the response of employment in columns (1) and (2), we note that the signs of the estimated coefficients are consistent with the model; focusing on the most saturated specification in column (2), the estimate of γ implies that an increase from the median to the 90th percentile in the exposure to labor-saving technologies implies a 4.4 percentage point decline in the cumulative growth of employment over the next decade. By contrast, moving from the median to the 90th percentile in exposure to labor-augmenting technologies implies a 9.1 percentage point increase in employment growth over the same period. On the other hand, columns (3) and (4) reveal that the average wage at the occupation level is essentially unrelated to our measure of labor-saving technology exposure, and only weakly related to our measure of labor-augmenting technology (average wages rise by 0.8 percentage points over the next decade following a shift from the median to 90th percentile in ξ^N). Consequently, the response of the total wage bill in columns (5) and (6) reflects primarily changes in employment.

Here, there are two points worth discussing. First, unlike the response of the earnings of incumbent workers, total employment, wages, and the total wage bill increase in response to improvements in labor-augmenting technology. This difference in individual outcomes versus outcomes for workers as a group is consistent with our model mechanism of skill displacement, in which the increased productivity of improved labor-augmenting technologies primarily benefits new workers rather than incumbents.

Second, the response of average wages in column (4) to labor-saving technologies reveals one difference between the model and the data. In the model, the response of individual earnings to an increase in labor-saving technology ξ^R in equation (16) is exactly equal to the change in average wages at the occupation level. In the data, individual earnings of incumbents fall following an increase in ξ^R as we saw in Table 1, while the response of average wages to ξ^R is essentially zero at the occupation level. Arguably, changes in employment status and worker composition play an important role in reconciling these estimates. In particular, the extensive margin of employment (which the model lacks) seems to play an important role in accounting for the earnings losses of individual workers, reinforcing our conclusions in Section 3.6. Reinforcing this point, recall that in Figure 4 we document that the most skilled workers were also most likely to experience involuntary job loss in response to an increase in ξ^N. By contrast, an increase in ξ^R led to an increase likelihood of job loss among the least skilled workers. These changes in worker composition tend to attenuate both the positive wage response to ξ^N and the negative wage response on ξ^R at the aggregated occupation-industry level, since now the worker pool is better following an increase in ξ^R, and worse following an increase in ξ^N. The fact that changes in worker composition can lead to these differences further emphasizes the value of studying the earnings growth of individual workers if the goal is to understand the displacive effects of technology.

4.2 Industry productivity and labor share

The spillovers of innovation across all workers—the last term in equation (8)—are a function of the growth in industry productivity, which is equal to the fall in the unit cost of production. It evolves according to

$$\begin{aligned}\Delta \log X &\approx \frac{1}{1+\chi\,\epsilon_c}\sum_{j\in J} s^p(j)\,\Gamma_j\,\varepsilon(j) \\ &= \frac{\Gamma_R}{1+\chi\,\epsilon_c}\,\frac{LS}{LS_R}\,\bar{\xi}^R + \frac{\Gamma_N}{1+\chi\,\epsilon_c}\,\frac{LS}{LS_N}\,\bar{\xi}^N,\end{aligned} \tag{20}$$

were ϵ_c denotes the elasticity of the marginal cost of output c_y to an increase in output, and $s^p(j)$ denotes the expenditure share of task j. Productivity growth in the model is a weighted average of task-specific technological improvements, weighted by the expenditure share of each task times the elasticity of its marginal cost to improvements in technology. The second line of equation (20) decomposes these technology improvements into improvements in labor-saving and labor-augmenting technologies at the industry level, where

$$\bar{\xi}^R \equiv \sum_o s(o)\xi^R(o), \quad \bar{\xi}^N \equiv \sum_o s(o)\xi^N(o). \tag{21}$$

These aggregated exposure measures (21) are weighted averages of the occupation-level exposures (12) weighted by the share $s(o)$ of occupation o in the industry wage bill.

We construct direct empirical analogues of (21) in the data,

$$\bar{\xi}^R_{k,t} \equiv \left[\sum_o s_{o,k,t}\,\xi^R_{o,k,t}\right], \qquad \bar{\xi}^N_{k,t} \equiv \left[\sum_o s_{o,k,t}\,\xi^N_{o,k,t}\right], \tag{22}$$

where we weigh our occupation-level exposure measures by the employment share $s_{o,k,t}$ of occupation o in industry k at year t. We then estimate

$$\frac{1}{h}\Big(\log(X_{k,t+h}) - \log(X_{k,t})\Big) = \gamma\,\bar{\xi}^R_{k,t} + \delta\,\bar{\xi}^N_{k,t} + c\,\mathbf{Z}_{k,t} + u_{k,t} \tag{23}$$

where the left hand side is the annualized growth in multi-factor productivity from the BLS at the NAICS4 industry level (see Appendix B.5 for details). To ensure that our results are not affected by broad industry trends, we include the interaction of coarse industry (NAICS2) fixed effects interacted by calendar year. We also include controls for the log of industry employment and lagged 3-year growth rates in the dependent variable. We weight regressions by yearly industry employment shares and focus on horizons of five years.

Column (1) of Table 4 reports the estimated coefficients γ and δ. We see that they are both positive and comparable in magnitude: a one standard deviation increase in either measure of technology exposure leads to approximately a 2.5 to 3.0 percentage point increase in the cumulative growth of industry productivity over the next five years. However, these coefficients are somewhat imprecisely estimated, partly because the two aggregated technology exposures are highly correlated (approximately 80%). Given that the estimated coefficients imply that differences between Γ_R and Γ_N together with LS_R and LS_N are likely small, we also estimate a version of equation (23) where we combine the two exposure measures into one, $\bar{\xi}_{k,t} = \bar{\xi}^R_{k,t} + \bar{\xi}^N_{k,t}$. The estimated coefficient in the combined exposure measure, reported in column (2), is much more precisely estimated and approximately equal to the sum of the two coefficients $\hat{\gamma}$ and $\hat{\delta}$, implying that the combined exposure measure summarizes all relevant information for industry productivity.

We next examine our model's predictions about the labor share. The industry labor share LS, defined as the ratio of the industry wage bill to the value of industry output, evolves according to

$$\begin{aligned}\Delta \log LS \approx \quad & \Gamma_R\left[(1-\nu_R)\frac{\psi+\zeta_R}{\nu_R+\zeta_R} + (\psi-1)\left(1-\frac{LS}{LS_R}\right) + \frac{LS}{LS_R}\frac{\vartheta}{1+\chi\epsilon_c}\right]\bar{\xi}^R \\ & + \Gamma_N\left[(1-\nu_N)\frac{\psi+\zeta_N}{\nu_N+\zeta_N} + (\psi-1)\left(1-\frac{LS}{LS_N}\right) + \frac{LS}{LS_N}\frac{\vartheta}{1+\chi\epsilon_c}\right]\bar{\xi}^N,\end{aligned} \tag{24}$$

where ϑ is the elasticity of the industry labor share to changes in productivity. In the model, the relation between the industry labor share and the overall level of overall labor-saving $\bar{\xi}_R$ and labor-augmenting $\bar{\xi}_N$ technologies is subtle, as it depends on three separate terms. The first term in each bracket captures a within-task reallocation effect: holding the task share of output constant, labor-saving technologies lower the industry labor share (assuming $\nu_R > 1$), while the effect of labor-augmenting technologies is positive if $\nu_N < 1$. The second term in brackets illustrate how technological innovation alters the allocation of expenditures across tasks, holding productivity fixed; their sign depends on the elasticity of substitution across tasks ψ and the relative labor intensity of routine and non-routine tasks. The last term in each bracket accounts for cross-task spillovers through changes in industry productivity and depends on the sign of ϑ.

We estimate the relation between our technology exposure measures and the industry labor share using equation (23). As we see in column (3) of Table 4, the response of the industry labor share is largely driven by the direct effects above. That is, an increase in exposure to labor-saving technologies leads to a 2.5 percent cumulative decline in the labor share over the next five years; by contrast, an increase in exposure to labor-augmenting technologies is followed by approximately a 0.75 percent increase, though the estimate is pretty imprecise and not statistically different from zero. In column (4) we show that the total wage bill at the industry level has a very similar response.

4.3 Spillovers to worker-level earnings

Having established a link between our technology exposure measures and industry productivity, we next turn our attention to estimating the spillover term (the last term in equation (8)). We do so using a modified version of (16),

$$\Delta w^i_{t+h} = \alpha\,\bar{\xi}_{i,t} + \gamma\,\xi^R_{i,t} + \delta\,\xi^N_{i,t} + c\,\mathbf{Z}_{i,t} + \varepsilon_{i,t}. \tag{25}$$

where now $\bar{\xi}_{k,t} = \bar{\xi}^R_{k,t} + \bar{\xi}^N_{k,t}$ is the combined technology exposure measure, which, given the discussion in the previous section, contains all the relevant information in predicting industry productivity. To identify our coefficient of interest α, we now replace the granular industry-year fixed effects with a coarser version (we move from 4-digit NAICS to 2-digits) and we now control for total industry employment (in logs). All other controls remain as in (16).

The estimated coefficient α in (25) is both economically and statistically significant. A one-standard deviation increase in $\bar{\xi}$ is associated with a 2.5 percentage point increase in the growth of (cumulative) worker earnings; the estimates of γ and δ from (25) are very close to our baseline estimates from (16). As we see in Appendix Table A.6, allowing the coefficients γ and δ to vary by prior income, or using our alternative technology exposure measure leads to quantitatively similar estimates for α.

4.4 Reconciling the estimates from the aggregate and micro-level regressions

Are the different regression estimates from the aggregate and micro data quantitatively consistent with the model? To answer this question, we next use these coefficient estimates to estimate the underlying parameters of the model using GMM. To do so, however, we need to make some additional assumptions on how the empirical measures relate to their model equivalents: we assume that our empirical measure of technology exposure (15) maps directly into its model counterpart, up to the presence of industry-level measurement error—capturing, for example, differences in patenting rates across industries. In brief, we target the coefficient estimates corresponding to our worker level regressions (16), conditioning on income; the response of employment at the occupation level to our exposure measures (19); and the response of productivity, labor share, and individual worker wages to our industry-level exposure measure (equations (23) and (25)). The model is over-identified: we have 16 moments plus two cross-equation restrictions to estimate 11 parameters. Table 5 reports the moments we target in our estimation and the resulting parameter estimates. Appendix A.8 contains all details of the estimation procedure.

The top panel of Table 5 shows that the overall fit of the model is rather good. Examining the parameters in the bottom panel, we see that they are largely consistent with the literature. Our estimate of ψ is close to one, in line with the estimates in Caunedo et al. (2023). The estimate of χ is close to 2, which is consistent within the range of estimates reported by Broda and Weinstein (2006).

The estimates of $\nu_R = 1.44$ and $\nu_N = 0.93$ are consistent with technology substituting routine and complementing non-routine tasks. The estimates of labor supply elasticities ζ_R and ζ_N are in line with existing estimates of the elasticity of labor supply based on micro data (Chetty, Guren, Manoli, and Weber, 2011). Last, our estimates of skill loss are modest: the decline in incumbent worker earnings in response to ν_N identifies the mean skill loss among incumbent workers to equal $\beta = 0.052$, while worker skill is reasonably persistent across vintages ($\omega = 0.014$).

5 Application to Artificial Intelligence

Recent advances in generative artificial intelligence have spurred fears about the impact of artificial intelligence (AI) on workers. Just like most other technologies, AI can affect worker earnings either because it automates certain tasks that workers do, or because it serves to complement their tasks—possibly at the cost of some skill displacement on the part of the worker. Here, we use our framework to make predictions about the impact of AI on worker earnings over the medium run.

We take an approach similar to our validation exercise in Section 2.2 to estimate the exposure of each task to AI. We first query GTP4 on whether AI could perform the same task with little to no human assistance, or whether it would be unable to perform the task on its own without significant human intervention. We group tasks that can be performed by AI with little human assistance as those that are substituted by AI, and we assign the remaining tasks as AI complements. Next, we query GPT4 on whether AI applications to these tasks are feasible now or in the near future. The answer to the second query maps into whether capital specific to task j experiences a quality improvement $\varepsilon_{AI} > 0$ in equation (4). Appendix A.9 contains all details.

Second, we use the model equation (8), together with our parameter estimates in Table 5, to construct predictions for the changes in worker earnings at the occupation level over the next decade using the 2021 ACS. When doing so we focus on only the first terms; the third term averages out to zero at the occupation level while is difficult to make predictions about cross-industry differences in adoption of AI and therefore infer productivity increases at the industry level. Thus, our predictions concern only the direct effects on AI on worker earnings. Last, we construct direct analogues of occupation-level exposures (12) that are specific to AI. To calibrate the size of the improvement ε_{AI} for affected tasks, we rely on the estimates of the model parameters in Table 5 to target a 10% decline in earnings (not considering skill displacement) for the worst-affected occupation over the next 5 years; this estimate corresponds to a recent example of earnings losses due to automation: following the introduction of e-commerce in the late 1990s, order clerks experienced a 20 percent decline in real earnings relative to other clerking occupations over the next decade.

Table 6 and Figure A.7 summarize the exposure of worker earnings to AI at a broad level. Comparing Figure A.7 to Figure 1, we see that the types of occupations that are exposed to the labor-saving effect of AI include not only production and transportation workers, but also occupations in the office and administrative support category and sales. In terms of its labor-augmenting applications, business and management occupations continue to be as highly exposed as past technologies, but now office and administrative workers are quite highly exposed as well. Examining Table 6, we see that the most adversely affected workers are in occupations that fall in the office and administrative category. Incumbents workers in these occupations, which comprise approximately 10 percent of overall employment, are predicted to experience relative declines in earnings of approximately 8.7 percent over the medium run. This earnings decline can be decomposed into a 6.9 percent decline due to AI automating their tasks; a 6.8 percent increase stemming from productivity improvements; and a 8.6 percent decline due to skill displacement.

We report the most exposed occupations to AI, in terms of automation or potential increases in productivity in Appendix Tables A.8 and A.9, respectively. In the list of most highly exposed occupations to AI in terms of automation are Tellers, Word Processors and Typists, Data Entry Keyers, and Bookkeeping, Accounting, and Auditing Clerks. Our prediction is that these occupations will experience significant declines in earnings over the next five years relative to other workers. The occupations most highly exposed to AI as a complementary technology are markedly different: Insurance Underwriters, Medical Transcriptionists, Customer Service Representatives, Personal Financial Advisors, and Budget Analysts. However, even though AI can help these workers perform these tasks more efficiently, incumbent workers may lack the skills to use it effectively. Overall, our model estimates imply that incumbent workers in these occupations are likely to experience relative declines in earnings of 3.0 to 7.8 percent over the next five years.

Here, a few caveats are in order. First, overall magnitudes are hard to assess, so we have more confidence in the relative ranking of different occupations than the level of the predicted growth rates. Second, and most importantly, these projections exclude the effect of spillovers arising from productivity improvements. Third, AI is likely the first vintage of the technology, hence the skill displacement effects could be significantly more muted than our estimates suggest; that said, Brynjolfsson, Li, and Raymond (2023) document that the introduction of a generative AI tool led to productivity increases only among low-skilled workers, consistent with the view that skills are not perfectly transferable across technologies. Last, differences in adoption rates across industries could generate additional heterogeneity not captured by our estimates.

6 Conclusion

We develop a methodology for identifying the arrival of labor-saving and labor-augmenting technologies that relies only on the textual description of the patent document and the tasks performed by workers in an occupation. Combining our measures with administrative data on worker earnings allows us to trace the impact of these technologies on the cross-section of workers. Overall, we find that labor-saving technologies have a uniformly negative impact on worker earnings and employment.

By contrast, labor-augmenting technologies have much more heterogenous impact on workers: they are followed by a modestly negative impact on the earnings of incumbent workers, though these declines are primarily concentrated in the most skilled workers, consistent with skill displacement. In addition, we see that labor-augmenting technologies lead to higher wages and employment at the occupation level, suggesting they primarily benefit new workers. Applying our calibrated model to the case of Artificial Intelligence (AI) allows us to make predictions about its impact on worker earnings over the medium run.

References

Abel, A. B. (2019). Classical measurement error with several regressors. Working paper, University of Pennsylvania.

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of Labor Economics, Volume 4. Amsterdam: Elsevier-North*, pp. 1043–1171.

Acemoglu, D. and P. Restrepo (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review 108*(6), 1488–1542.

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy 128*(6), 2188–2244.

Acemoglu, D. and P. Restrepo (2022). Tasks, automation, and the rise in u.s. wage inequality. *Econometrica 90*(5), 1973–2016.

Aghion, P., C. Antonin, S. Bunel, and X. Jaravel (2021). What are the labor and product market effects of automation? new evidence from france. Working paper.

Akerman, A., I. Gaarder, and M. Mogstad (2015). The skill complementarity of broadband internet. *The Quarterly Journal of Economics 130*(4), 1781–1824.

Arora, S., Y. Liang, and T. Ma (2017). A simple but tough-to-beat baseline for sentence embeddings. In *ICLR*.

Atack, J., R. A. Margo, and P. W. Rhode (2019). Automation of manufacturing in the late nineteenth century: The hand and machine labor study. *Journal of Economic Perspectives 33*(2), 51–70.

Atack, J., R. A. Margo, and P. W. Rhode (2022). "Mechanization takes command?": Powered machinery and production times in late nineteenth-century american manufacturing. *The Journal of Economic History 82*(3), 663?689.

Autor, D., C. Chin, A. Salomons, and B. Seegmiller (2022). New frontiers: The origins and content of new work, 1940–2018. Working paper, MIT.

Autor, D., D. Dorn, G. H. Hanson, and J. Song (2014). Trade adjustment: Worker-level evidence. *The Quarterly Journal of Economics 129*(4), 1799–1860.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5), 1553–97.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006). The Polarization of the U.S. Labor Market. *American Economic Review 96*(2), 189–194.

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

Bárány, Z. L. and C. Siegel (2018). Job polarization and structural change. *American Economic Journal: Macroeconomics 10*(1), 57–89.

Benmelech, E. and M. Zator (2022). Robots and firm investment. Working Paper 29676, National Bureau of Economic Research.

Bessen, J., M. Goos, A. Salomons, and W. van den Berge (2023). What Happens to Workers at Firms that Automate? *The Review of Economics and Statistics*, 1–45.

Bonfiglioli, A., R. Crinò, H. Fadinger, and G. Gancia (2020). Robot Imports and Firm-Level Outcomes. Crc tr 224 discussion paper series, University of Bonn and University of Mannheim, Germany.

Braxton, J. C., K. F. Herkenhoff, J. L. Rothbaum, and L. Schmidt (2021). Changing income risk across the US skill distribution: Evidence from a generalized kalman filter. Technical report, National Bureau of Economic Research.

Braxton, J. C. and B. Taska (2020). Technological change and the consequences of job loss.

Broda, C. and D. E. Weinstein (2006). Globalization and the Gains From Variety*. *The Quarterly Journal of Economics 121*(2), 541–585.

Brynjolfsson, E., D. Li, and L. R. Raymond (2023). Generative ai at work. Working Paper 31161, National Bureau of Economic Research.

Caunedo, J., D. Jaume, and E. Keller (2023). Occupational exposure to capital-embodied technical change. *American Economic Review 113*(6), 1642–85.

Chari, V. V. and H. Hopenhayn (1991). Vintage human capital, growth, and the diffusion of new technology. *Journal of Political Economy 99*(6), 1142–1165.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review 101*(3), 471–75.

Dauth, W., S. Findeisen, J. Suedekum, and N. Woessner (2021). The Adjustment of Labor Markets to Robots. *Journal of the European Economic Association 19*(6), 3104–3153.

Dechezleprêtre, A., D. Hémous, M. Olsen, and C. Zanella (2021). Induced automation: evidence from firm-level patent data. *University of Zurich, Department of Economics, Working Paper* (384).

Deming, D. J. and K. Noray (2020). Earnings dynamics, changing job skills, and stem careers. *The Quarterly Journal of Economics 135*(4), 1965–2005.

Feigenbaum, J. and D. P. Gross (2020). Automation and the fate of young workers: Evidence from telephone operation in the early 20th century. Working Paper 28061, National Bureau of Economic Research.

Fort, T. C. and S. Klimek (2018). The effects of industry classification changes on us employment composition. Working Paper 18-28.

Goldin, C. and L. Katz (2008). *The Race Between Education and Technology*. Harvard University Press.

Goldin, C. and L. F. Katz (1998). The origins of technology-skill complementarity. *The Quarterly Journal of Economics 113*(3), 693–732.

Goldschlag, N., T. J. Lybbert, and N. J. Zolas (2016). An 'algorithmic links with probabilities' crosswalk for uspc and cpc patent classifications with an application towards industrial technology composition. CES Discussion Paper 16-15, U.S. Census Bureau.

Goos, M. and A. Manning (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The Review of Economics and Statistics 89*(1), 118–133.

Gouin-Bonenfant, E. (2020). Productivity dispersion, between-firm competition, and the labor share.

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

Guvenen, F., S. Ozkan, and J. Song (2014). The Nature of Countercyclical Income Risk. *Journal of Political Economy 122*(3), 621–660.

Haltiwanger, J., R. Jarmin, R. Kulick, J. Miranda, and V. Penciakova (2019). Augmenting the lbd with firm-level revenue. Technical Report CES-TN-2019-02, US Census Bureau.

Hemous, D. and M. Olsen (2021). The rise of the machines: Automation, horizontal innovation, and income inequality. *American Economic Journal: Macroeconomics*.

Hicks, J. (1932). *The Theory of Wages*. Macmillan.

Hombert, J. and A. Matray (2021). Technology boom, labor reallocation, and human capital depreciation. *Working Paper*.

Hornstein, A., P. Krusell, and G. L. Violante (2005). The Effects of Technical Change on Labor Market Inequalities. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 1 of *Handbook of Economic Growth*, Chapter 20, pp. 1275–1370. Elsevier.

Hornstein, A., P. Krusell, and G. L. Violante (2007). Technology—Policy Interaction in Frictional Labour-Markets. *Review of Economic Studies 74*(4), 1089–1124.

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

Jarmin, R. S. and J. Miranda (2002). The longitudinal business database. Technical report.

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

Karabarbounis, L. and B. Neiman (2013). The Global Decline of the Labor Share*. *The Quarterly Journal of Economics 129*(1), 61–103.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2021). Measuring technological innovation over the long run. *American Economic Review: Insights 3*(3), 303–20.

Kerr, W. and S. Fu (2008). The survey of industrial r&d—patent database link project. *The Journal of Technology Transfer 33*, 173–186.

Koch, M., I. Manuylov, and M. Smolka (2021). Robots and Firms. *The Economic Journal 131*(638), 2553–2584.

Kogan, L., D. Papanikolaou, L. D. W. Schmidt, and B. Seegmiller (2021). Technology, vintage-specific human capital, and labor displacement: Evidence from linking patents with occupations. Working Paper 29552, National Bureau of Economic Research.

Koh, D., R. Santaeulalia-Llopis, and Y. Zheng (2020). Labor share decline and intellectual property products capital. *Econometrica 88*(6), 2609–2628.

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*(5), 1029–1053.

Mann, K. and L. Püttmann (2018). Benign effects of automation: New evidence from patent texts. *Working Paper*.

Mikolov, T., I. Sutskever, K. Chen, G. Corrado, and J. Dean (2013). Distributed representations of words and phrases and their compositionality. *CoRR abs/1310.4546*.

Pennington, J., R. Socher, and C. D. Manning (2014). Glove: Global vectors for word representation. In *EMNLP*.

Seegmiller, B. (2021). Valuing labor market power: The role of productivity advantages. Working paper, MIT.

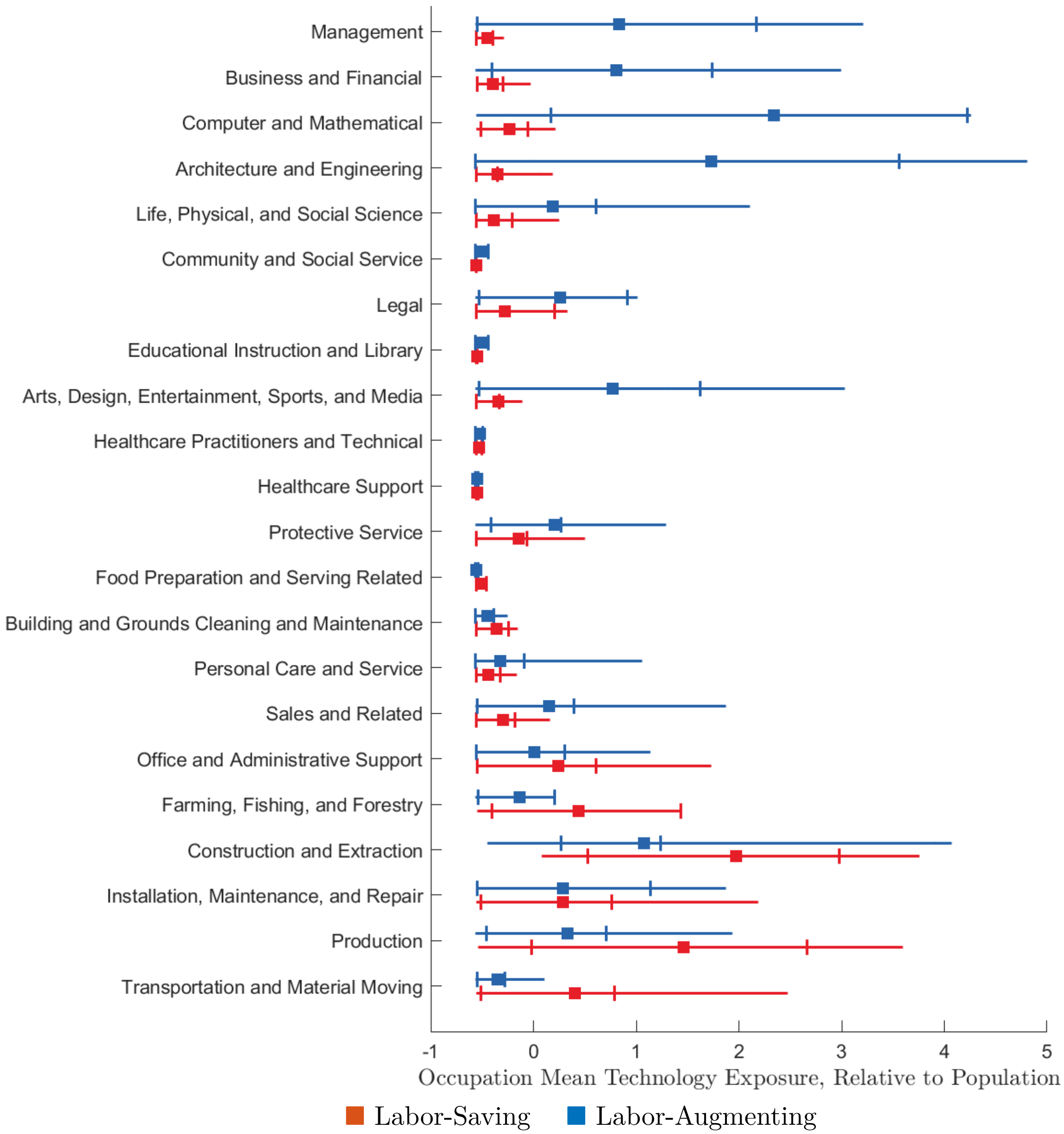
Seegmiller, B., D. Papanikolaou, and L. D. Schmidt (2023). Measuring document similarity with weighted averages of word embeddings. *Explorations in Economic History 87*(C), S0014498322000729.

Violante, G. L. (2002). Technological acceleration, skill transferability, and the rise in residual inequality. *The Quarterly Journal of Economics 117*(1), 297.

Webb, M. (2020). The impact of artificial intelligence on the labor market.

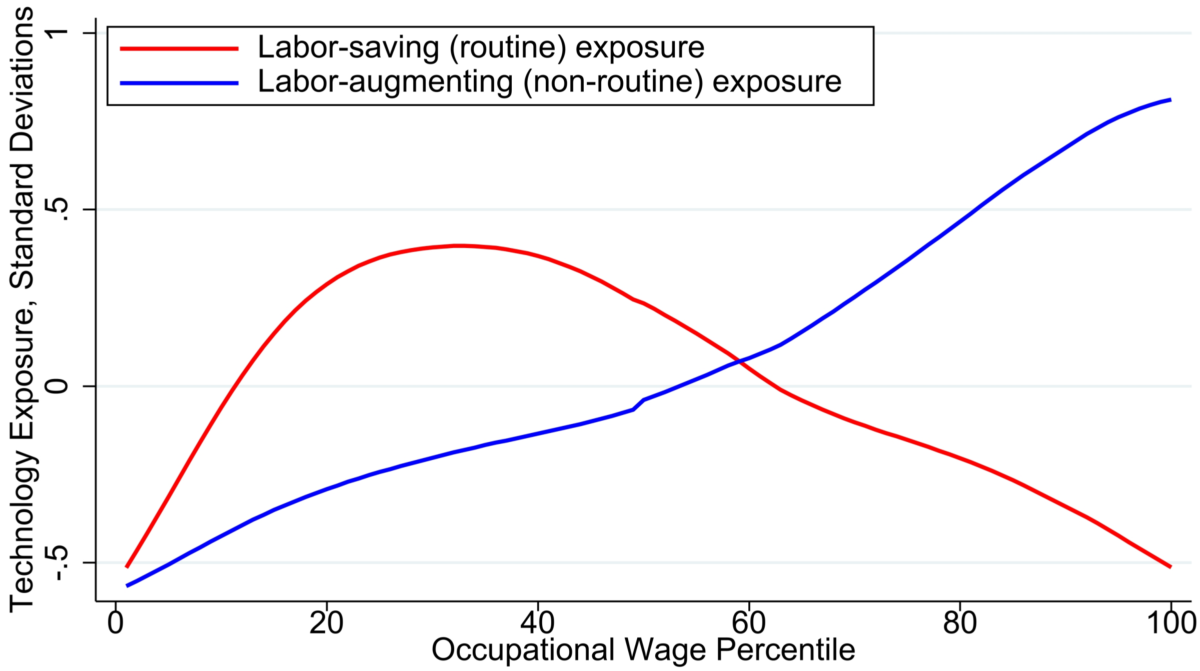
Figures and Tables

Figure 1: Distribution of technology exposure across occupations



Note: The figure shows the cross-sectional distribution of the measure of industry–occupation technology exposure to labor-saving and labor-augmenting technologies by broad occupation category. The square corresponds to the average exposure within each broad group; the vertical marks correspond to the 25th and 75th percentile of the average exposure of occupations within each broad occupation category; the length of the lines corresponds to the p90-p10 range. Both routine and non-routine exposures are first standardized to zero mean and unit standard deviation level at the employment-weighted population level. Broad occupation categories correspond to 2-digit SOC codes; occupations correspond to 6-digit SOC codes and industries at the 4-digit NAICS level.

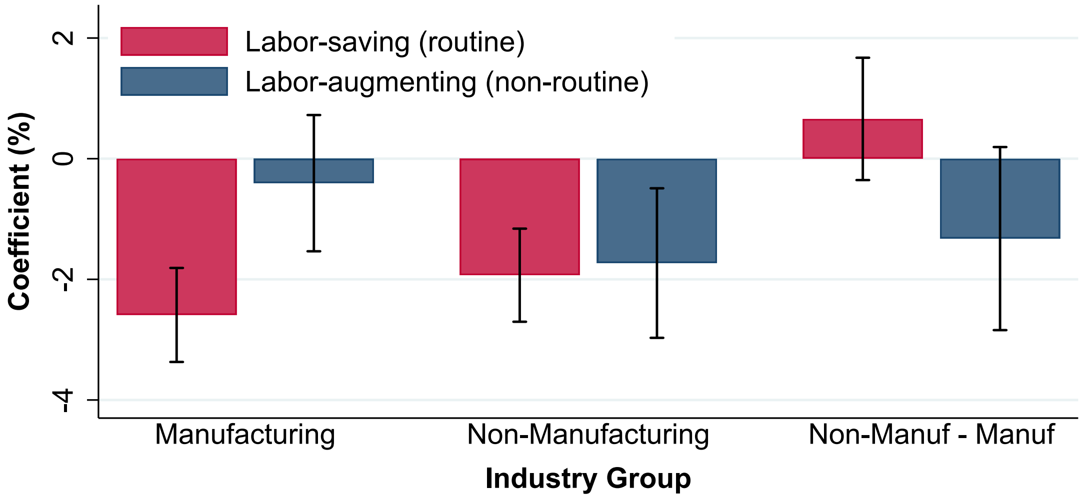
Figure 2: Technology exposure by occupation skill (wage)



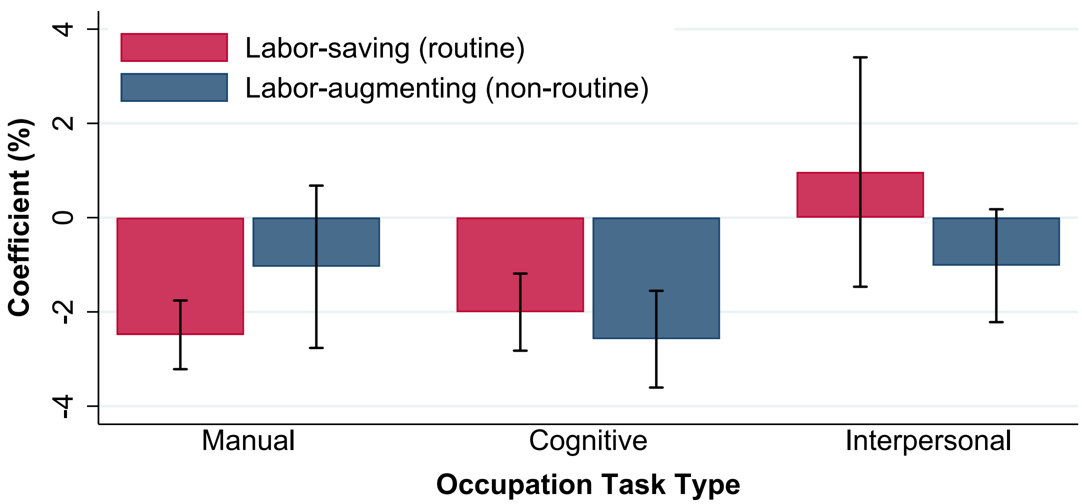
Note: This figure shows average routine or non-routine technological exposure (given by (15) in the main text) plotted against occupational cross-sectional wage percentile rank (lowess smoothed using a bandwidth of 5 bins). We compute yearly averages of technological exposure in employment-weighted standard deviation units at the occupation level and then rank occupations into percentiles based on their average wages in a given year. We generate this figure using a version of our measure that is computed using publicly available sources based on patent-to-industry probabilistic links from Goldschlag, Lybbert, and Zolas (2016), as well as Decennial Census data on occupational employment and wages for the 1980, 1990, and 2000 Census years. See main text and appendix section B.6 for further details.

Figure 3: Technology exposure and worker earnings growth, by industry or occupation type

A. Manufacturing vs Non-Manufacturing

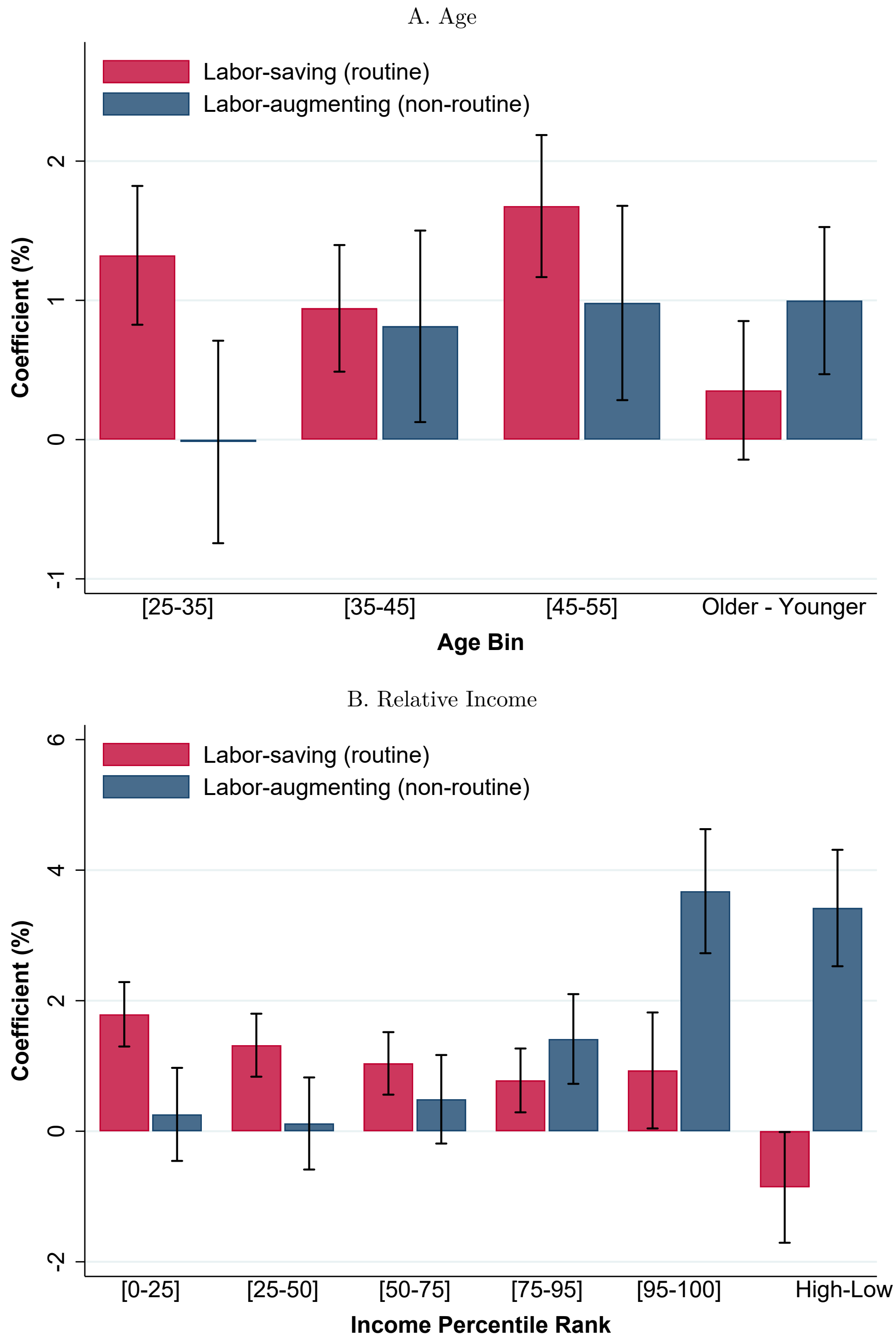


B. Occupation Task Type: Manual/Cognitive/Interpersonal



Note: This figure shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text; we allow these coefficients to vary by industry type (panel A) or occupation task type (panel B). In panel A we compare coefficients for individuals employed in or out of manufacturing (broadly defined as 2-digit NAICS codes 11 through 33). Our broad definition of manufacturing also includes construction, mining/extraction, utilities, and agriculture, but a large majority of total employment from this wider set of industries still comes from the standard NAICS manufacturing industries 31 to 33. In panel B we designate occupations as primarily focusing on either manual, cognitive, or interpersonal tasks using task scores from Acemoglu and Autor (2011). The Acemoglu and Autor (2011) non-routine manual (physical) and routine manual task category scores gives the manual score; average non-routine cognitive (analytical) and routine cognitive represents the cognitive score; and non-routine manual (interpersonal) and non-routine cognitive (interpersonal) scores yields the interpersonal score. Each year we assign occupations into one of these three disjoint categories depending on which of the three task scores corresponds to the highest cross-sectional percentile ranking for that occupation. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We plot 95% confidence intervals from standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. All specifications include industry × year, occupation × year, within occupation–industry income bin × year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Figure 4: Technology exposure and worker job loss, by worker age or relative income



Note: This figure shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text, where we replace earnings growth with a proxy for involuntary job loss as the dependent variable; we allow these coefficients to vary by age group (panel A) or within occupation-industry income rank (panel B). The dependent variable is an indicator for whether or not a worker leaves their current employer in the next five years and also experiences an income growth rate beneath the 20th percentile for the year. We plot 95% confidence intervals from standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. All specifications include industry $\times$ year, occupation $\times$ year, within occupation–industry income bin $\times$ year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Table 1: Technology exposure and worker earnings growth

	(1)	(2)	(3)	(4)
A. Horizon: 3 years of worker earnings growth				
Labor-saving Exposure ξ^R	-1.987	-1.960	-2.056	-2.156
	(-9.53)	(-9.31)	(-7.71)	(-7.75)
Labor-augmenting Exposure ξ^N	-0.764	-0.619	-0.868	-0.974
	(-2.82)	(-2.29)	(-2.15)	(-2.30)
B. Horizon: 5 years of worker earnings growth				
Labor-saving Exposure ξ^R	-2.168	-2.049	-2.371	-2.428
	(-9.15)	(-8.62)	(-7.97)	(-7.83)
Labor-augmenting Exposure ξ^N	-0.773	-0.573	-1.182	-1.296
	(-2.52)	(-1.9)	(-2.59)	(-2.72)
C. Horizon: 10 years of worker earnings growth				
Labor-saving Exposure ξ^R	-2.559	-2.315	-2.850	-2.854
	(-9.00)	(-8.18)	(-7.90)	(-7.67)
Labor-augmenting Exposure ξ^N	-0.828	-0.625	-1.515	-1.678
	(-2.37)	(-1.79)	(-2.74)	(-2.98)
Fixed Effects				
Industry (NAICS 4-digit)	Y	Y		
Occupation (occ1990dd)	Y		Y	
Ind × Year			Y	Y
Occ × Year		Y		Y
Prior Income Rank × Year	Y	Y	Y	Y

Note: Table shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 3, 5, or 10 years. We report t statistics corresponding to standard errors clustered at the occupation-industry (NAICS 4-digit) level in parentheses beneath the coefficient estimates. We scale the coefficients to that they correspond to a shift from the median to the 90th percentile of our technology exposure measures. The vector of controls $\mathbf{Z}$ includes flexible non-parametric controls for worker age and the level of past worker earnings as well as recent earnings growth rates. In addition, we include different combinations of year, occupation, worker prior income, and industry fixed effects, depending on the specification. We construct controls for worker age and lagged earnings by linearly interpolating between 3rd degree Chebyshev polynomials in workers' lagged income quantiles within an industry-age bin at 10-year age intervals. In addition, to soak up some potential variation related to potential mean-reversion in earnings (which could be the case following large transitory shocks), we also include 3rd degree Chebyshev polynomials in workers' lagged income growth rate percentiles, and we allow these coefficients to differ by gender as well as past income levels based on five gender-specific bins formed based upon a worker's rank relative to her peers in the same industry and occupation. Columns (1) to (4) report coefficient estimates under different combinations of fixed effects; all specifications include prior income rank × calendar year fixed effects. Prior income rank bins are based on workers' yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there are fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification when assigning income ranks. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th and 95th percentile; and, 95th percentile and above.

Table 2: Worker technology exposure and earnings growth, by skill

	Technology Exposure	
	Labor-Saving (ξ^R)	Labor-Augmenting (ξ^N)
A. Worker Age		
25-35 y/o	-2.62	-0.08
	(-7.32)	(-0.14)
35-45 y/o	-1.84	-1.57
	(-5.44)	(-3.12)
45-55 y/o	-3.16	-1.80
	(-8.26)	(-3.52)
Older - Younger	-0.54	-1.72
	(-1.51)	(-4.25)
B. Income (relative to Ind × Occ peers)		
0–25th percentile	-2.72	-0.82
	(-7.16)	(-1.49)
25–50th percentile	-1.96	-0.80
	(-5.49)	(-1.53)
50–75th percentile	-2.30	-1.11
	(-6.86)	(-2.25)
75–95th percentile	-2.58	-2.09
	(-7.07)	(-3.93)
95–100th percentile	-3.12	-4.74
	(-4.89)	(-5.66)
Top - Bottom	-0.39	-3.92
	(-0.62)	(-4.49)
C. Education		
No College Education	-2.38	-1.03
	(-7.72)	(-2.18)
College Educated	-2.82	-1.58
	(-5.67)	(-3.18)
College - No College	-0.44	-0.55
	(-1.02)	(-2.12)

Note: This table shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text; we allow these coefficients to vary by age group (panel A); within occupation-industry income rank (panel B); or 4-year college graduate status (panel C). The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report t-statistics based on standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. All specifications include industry × year, occupation × year, within occupation–industry income bin × year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Table 3: Technology exposure and aggregate worker outcomes
(occupation–industry level)

	Employment		Avg Wage		Wage Bill	
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure to labor-saving (ξ^R)	-0.241	-0.440	0.000	-0.003	-0.240	-0.463
	(-1.93)	(-3.52)	(0.07)	(-1.13)	(-1.92)	(-3.69)
Exposure to labor-augmenting (ξ^N)	1.412	0.908	0.000	0.008	1.417	1.006
	(9.05)	(5.59)	(-0.05)	(2.04)	(9.07)	(6.13)
Observations	65,500	65,500	65,500	65,500	65,500	65,500
Fixed Effects						
Occupation × year FE	Y	Y	Y	Y	Y	Y
NAICS4 × year FE	Y	Y	Y	Y	Y	Y
Demographic controls		Y		Y		Y
Lagged Employment		Y		Y		Y
Lagged Wage		Y		Y		Y

Note: This table shows the estimated slope coefficients γ and δ (times 100) from equation (19) in the main text. The dependent variables are the annualized log changes in employment (columns 1 and 2); average wage (columns 3 and 4); or wagebill over the next 10 years for an aggregate occupation-industry cell. We report t-statistics based on standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. The data come from the 1980, 1990, and 2000 Decennial Census, as well as the 2008-2012 5-year ACS panel. We aggregate occupational employment, wages, wagebill, and demographic characteristics by occupation code and industry cells. For this analysis we create a modified version of 4-digit NAICS codes that allows us to construct a consistent crosswalk between the Census industry codes found in the Decennial/ACS and modified NAICS codes, and we re-compute all measures analogously at the modified NAICS industry level. We use restricted-access versions of these surveys available on Census data servers; the main advantage of this version of the data is that earnings survey responses are not top-coded, as is the case with the publicly available versions. Observation counts are rounded in accordance with Census disclosure rules. See appendix B.7 for details on the construction of the aggregated occupation-industry-level panel.

Table 4: Technology exposure and industry outcomes
(NAICS4 level)

	Productivity		Labor share	Wage Bill
	(1)	(2)	(3)	(4)
Overall Technology Exposure ($\bar{\xi}$)		1.085		
		(6.09)		
Exposure to labor-saving ($\bar{\xi}^R$)	0.504		-0.510	-0.626
	(1.67)		(-3.37)	(-2.88)
Exposure to labor-augmenting ($\bar{\xi}^N$)	0.628		0.141	0.291
	(3.01)		(1.31)	(1.49)
NAICS2 x year FE	y	y	y	y
Log Industry Employment	y	y	y	y
Lagged 3-year growth rate	y	y	y	y
Observations	2,700	2,700	2,700	2,700

Note: This table shows the estimated slope coefficients γ and δ (times 100) from equation (23) in the main text. The dependent variables are the annualized log changes in industry productivity index (columns 1 and 2); labor share (columns 3 and 4); or wagebill over the next 5 years. We report t-statistics based on standard errors clustered at the industry level beneath coefficient estimates; the coefficients correspond to a standard deviation increases in the given measure. Regressions additionally control for the log of industry employment and lagged 3-year growth rates in the dependent variable. Industry-level measures are yearly the employment-weighted averages of the worker-level occupation-industry exposures as in equation 15. The overall technology exposure is the standardized average of the two individual technology exposures. Data on industry outcomes are from the Bureau of Labor Statistics and our sample period covers 1987-2012. Observation counts are rounded in accordance with Census disclosure rules. See appendix B.5 for details on the construction of the industry-level outcomes from the BLS data.

Table 5: Model estimates

A. Target Moments		
Moment or regression coefficient in (%)	Model	Data
Response of worker earnings to ξ^R, homogeneous coefficient	-2.71	-2.43
Response of worker earnings to ξ^N, homogeneous coefficient	-1.38	-1.30
Response of worker earnings to ξ^N, 0–25th percentile in relative income	-0.19	-0.82
Response of worker earnings to ξ^N, 25–50th percentile in relative income	-1.08	-0.80
Response of worker earnings to ξ^N, 50–75th percentile in relative income	-1.69	-1.11
Response of worker earnings to ξ^N, 75–95th percentile in relative income	-2.39	-2.09
Response of worker earnings to ξ^N, 95–100th percentile in relative income	-3.32	-4.74
Response of worker earnings to ξ^R (regression with $\bar{\xi}$), homogeneous coefficient	-2.56	-2.99
Response of worker earnings to ξ^N (regression with $\bar{\xi}$), homogeneous coefficient	-1.50	-1.75
Response of worker earnings to $\bar{\xi}$, homogeneous coefficient	2.51	2.48
Response of industry productivity to composite exposure measure $\bar{\xi}$	5.61	5.43
Response of industry labor share to labor-saving tech exposure $\bar{\xi}^R$	-2.29	-2.55
Response of industry labor share to labor-augmenting tech exposure $\bar{\xi}^N$	0.63	0.71
Response of employment to labor-saving tech exposure ξ^R	-2.18	-2.20
Response of employment to labor-augmenting tech exposure ξ^N	4.64	4.54
Labor share (%), mean value	70.56	70.55

B. Parameter Estimates		
Parameter	Symbol	Estimate
Elasticity of substitution across tasks	ψ	1.191
Elasticity of substitution across industries	χ	1.910
Elasticity of substitution between capital and labor, routine tasks	ν_R	1.439
Elasticity of substitution between capital and labor, non-routine tasks	ν_N	0.928
Elasticity of labor supply, routine tasks	ζ_R	0.725
Elasticity of labor supply, non-routine tasks	ζ_N	0.871
Mean skill loss for incumbent workers	β	0.052
Capital-labor expenditure ratio, routine tasks	κ_R	0.349
Capital-labor expenditure ratio, non-routine tasks	κ_N	0.452
Skill loss across technology vintages	ω	0.014
Ratio of noise volatility to signal volatility for industry $\bar{\xi}^R$, $\bar{\xi}^N$ and $\bar{\xi}$	γ	1.035

Note: This table presents GMM estimates of our model from section 1 in the main text taken by minimizing the objective function (A.82) across 11 parameters (plus the two cross-equation restrictions $\Gamma_R = \Gamma_N$ and $A_R = A_N$) to match 16 target moments with their model-implied counterparts. We use a diagonal weighting matrix in (A.82) where weights correspond to the inverse of squared standard errors of empirical estimates. The empirical average labor share estimate and its standard error come from Koh, Santaeulalia-Llopis, and Zheng (2020). We calibrate coefficients to a 5-year horizon, and we scale the individual-level worker earnings responses to technology exposure to correspond to an increase from the median to 90th percentile of their empirical distributions, and we scale responses to industry-level aggregates $\bar{\xi}$, $\bar{\xi}^R$, and $\bar{\xi}^N$ to correspond to an empirical standard-deviation increase. See appendix section A for model details, including subsection A.8 for an in-depth discussion of model estimation. Also see Table A.7 for the list of closed-form model-implied coefficients in terms of model parameters.

Table 6: Occupational predicted earnings growth from exposure to AI, broad categories

Occupation	2-digit SOC	Automation	Complementary	Skill Displacement	Total	% of Emp
Management	11	-0.74	4.65	-5.84	-1.93	12.8
Business and Financial	13	-1.51	7.67	-9.63	-3.47	6.57
Computer and Mathematical	15	-1.98	5.82	-7.31	-3.47	4.12
Architecture and Engineering	17	-0.88	3.95	-4.96	-1.89	2.56
Science	19	-0.83	4.24	-5.32	-1.91	1.25
Community and Social Service	21	-0.29	2.52	-3.16	-0.93	1.84
Legal	23	-0.77	3.28	-4.12	-1.61	1.33
Education and Library	25	-0.54	3.89	-4.88	-1.53	6.16
Arts, Entertainment, Media	27	-1.21	3.96	-4.97	-2.22	1.66
Healthcare Practitioners	29	-1.19	3.46	-4.34	-2.08	6.65
Healthcare Support	31	-1.54	4.71	-5.92	-2.74	3.03
Protective Service	33	-2.03	4.63	-5.82	-3.21	2.45
Food Preparation and Serving	35	-4.01	4.28	-5.37	-5.10	3.76
Cleaning and Maintenance	37	-3.17	3.59	-4.50	-4.08	2.77
Personal Care and Service	39	-1.62	2.75	-3.45	-2.32	1.46
Sales and Related	41	-3.79	8.11	-10.2	-5.86	8.59
Office and Administrative	43	-6.92	6.82	-8.57	-8.66	10.6
Farming, Fishing, and Forestry	45	-3.33	3.06	-3.84	-4.11	0.61
Construction and Extraction	47	-0.98	3.13	-3.93	-1.78	4.71
Installation and Repair	49	-1.42	2.73	-3.42	-2.12	3.44
Production	51	-5.47	3.43	-4.31	-6.34	5.85
Transportation	53	-5.58	5.09	-6.38	-6.88	7.73
Overall		-2.74	4.95	-6.22	-4.01	100

Note: In this table we use model parameter estimates and occupational AI technological exposure proxies to predict the contributions of different components of the earnings growth equation (8) on incumbent workers in the given occupation category due to artificial intelligence. We calibrate to a 5-year horizon. The automation component corresponds to the employment-weighted average of $\left[\frac{\psi-\nu_R}{\nu_R+\zeta_R}\,\Gamma_R\right]\xi^R_{AI}(i)$ within the broad occupation category, where the AI subscript denotes our proxy for AI exposure. The complementary component represents the average of $\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\,\Gamma_N\right]\xi^N_{AI}(i)$, and the displacement component corresponds to $-\beta\xi^N_{AI}(i)$. The total component is the sum of the three. Because it averages to zero within occupation, the heterogeneous displacement term $\omega\left[\log\bar{l}(i)-\int[\log\bar{l}(i)]\,dF(i)\right]\xi^N(i)$ plays no role, and we also abstract away from aggregate productivity improvements driven by $\Delta\log X$. Data on employment comes from the 2021 ACS. See main text and appendix A (especially subsection A.9) for further details on how we construct the AI exposure measures and use them to compute model-implied wage growth.

A Model Appendix

In this section, we provide additional details about the model derivation and calibration.

A.1 Setup

For convenience, we repeat some of the assumptions of the model from section 1 of the main text here.

Aggregate output is a CES aggregate over output produced across a continuum of industries

$$\bar{Y} = \left(\int_k Y(k)^{\frac{\chi-1}{\chi}} dk \right)^{\frac{\chi}{\chi-1}}, \tag{A.1}$$

where $\bar{Y}$ is a constant and $\chi > 0$ is a parameter which captures the elasticity of demand for industry output. Industries are competitive and there is free entry of all firms in the production sector. In equilibrium, prices of industry output are equal to its marginal cost, and given the constant returns to production firms make zero profits. As stated in the main text, we then focus on a given industry and drop the k subscripts unless needed.

The industry-level output Y is a CES aggregate of a large number of intermediate tasks, indexed by $j \in \{1, \dots, J\}$:

$$Y = \left(\sum_{i=1}^{J} y(j)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}}, \tag{A.2}$$

where the parameter $\psi > 0$ indexes the elasticities of substitution across tasks j, as well as the absolute value of the demand elasticity for each task output. Each task j is produced using $k(j)$ and labor $l(j)$ according to

$$y(j) = \left((1-\gamma_j) k(j)^{\frac{\nu_j-1}{\nu_j}} + \gamma_j l(j)^{\frac{\nu_j-1}{\nu_j}} \right)^{\frac{\nu_j}{\nu_j-1}}, \tag{A.3}$$

where the parameter $\nu_j > 0$ governs the elasticities of substitution between $l(j)$ and $k(j)$ when producing each task and the parameter $\gamma_j \in (0, 1)$. As stated in the main text, we partition the set of tasks into routine and non-routine tasks, so $J = J_R \cup J_N$, and $\nu_j \in \{\nu_R, \nu_N\}$ with $\nu_N < \nu_R$.

We model the cumulative impact of an industry's breakthrough innovations between the first and second period as lowering the price of capital specific to task j,

$$\Delta \log q(j) = -\varepsilon(j) \tag{A.4}$$

where $\varepsilon \equiv [\varepsilon_1 \dots \varepsilon_J]$ is a vector of weakly positive random variables jointly distributed according to $f(\varepsilon)$.

There is a continuum of measure I workers supplying labor across the different tasks j, so that the aggregate supply of labor in task j, $L(j)$, is given by:

$$L(j) = \int_0^I l(i,j) di, \tag{A.5}$$

where $l(i,j)$ is the number of efficiency units of labor supplied by worker i in task j.

Each worker i is associated with a single occupation $o(i)$. Workers in occupation o supply positive output in for a small number of routine tasks and non-routine tasks, we denote by $J_o \in J$ the set of tasks performed by occupation o, and we have $\cup J_o = J$ and $\cap J_o = \emptyset$.

There are changes in aggregate and individual labor supply after technology shocks. We allow for skill displacement at the individual level: the adoption of a new vintage of technology leads to a change not only in wages but also in the number of efficiency units of a worker's human capital. Specifically, changes in worker i's productivity in task j satisfies

$$\Delta \log l(i,j) = \begin{cases} -\beta I[i \in J_N]\varepsilon(j) + u_{i,j} - \log l(i,j) & \text{w/ prob. } I[i \in J_N]\,\omega\,\varepsilon(j) \\ -\beta I[i \in J_N]\varepsilon(j) & \text{otherwise} \end{cases}, \tag{A.6}$$

where ω is a positive constant indicating the strength of the skill displacement effect, and $u(i,j)$ a new, i.i.d. draw from the ergodic distribution of $\log z(i,j)$ for incumbent workers in the same occupation. The first term says that larger shifts in the technology frontier are likely to generate greater amounts of skill displacement among incumbent workers, while the second term is a force of redistribution: there is some mean reversion in workers' productivity levels, so the most productive workers in the current vintage are most likely to experience declines in their productivity if technology changes.

The aggregate labor supply for task j varies on the extensive margin. We assume that aggregate labor supply satisfies

$$\Delta \log L(j) = \bar{\zeta} + \zeta_j \Delta \log w(j), \tag{A.7}$$

where ζ_j indexes the elasticity of aggregate labor supply in performing task j to labor compensation and $\bar{\zeta}$ is a constant which plays no material role in our analysis. Equation (A.7) makes implicit assumptions about the labor supply of new workers. In particular, the quantity of labor supplied by incumbent workers is pinned down by equation (A.6). Changes in aggregate labor supply are also affected by the entry of new workers. Equation (A.7) and equation (A.6) imply that new entrants are, on average, more highly skilled in the new technology than incumbents – exactly so that the aggregate loss of skill for incumbent workers is offset by the higher skill of new entrants, implying the absence of skill displacement in the aggregate.

For further notational convenience, according to (A.7), we capture the labor supply function in each period

$$L(j) = L_0(j) w(j)^{\zeta(j)} \tag{A.8}$$

where $L_0(j)$ is a constant which captures impacts of $\bar{\zeta}$ and also potentially other shifters of the aggregate supply labor in task j. We treat $L_0(j)$ as fixed except in the extension of the model where we allow aggregate skill displacement in section A.6.

A.2 Equilibrium conditions

A competitive equilibrium in this economy is a set of prices and allocations which satisfies the following conditions:

- The representative firm chooses $l^*(j), k^*(j)$ to maximize profits, taking factor prices $w^*(j), q(j)$ and the output price c_y^* as given,
- Labor supply equals labor demand, $L(j) = l^*(j)$, and
- Output markets clear: the quantity demanded equals the quantity supplied given output price c_y^* (which equals the marginal cost of production), $l^*(j)$, and $k^*(j)$.

Consider the firm's cost minimization problem across tasks j:

$$\min_{y(j)} \sum_{j \in J} p(j) y(j) \qquad s.t. \;\; Y = \left(\sum_{j \in J} y(j)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}} \tag{A.9}$$

where $y(j) = \left[(1-\gamma_j)k(j)^{\frac{\nu_j-1}{\nu_j}} + \gamma_j\, l(j)^{\frac{\nu_j-1}{\nu_j}}\right]^{\frac{\nu_j}{\nu_j-1}}$. Here $p(j)$ is the per-unit cost index for task j after optimal input choices have been made within task j. Under perfect competition, $p(j)$ equals the firm's per-unit marginal cost of producing $y(j)$. Our CES structure admits the following Hicksian demand for $y(j)$ from the first order condition for problem (A.9):

$$y(j) = \frac{1}{p(j)^{\psi}}\left[\sum_{j\in J} p(j)^{1-\psi}\right]^{\frac{\psi}{1-\psi}} Y = p(j)^{-\psi} X^{\chi-\psi}\bar{Y}, \tag{A.10}$$

where the second equality in (A.10) holds from defining the level of industry productivity (output per dollar of input expediture) X as

$$X \equiv \left(\sum_{j=1}^{J} p(j)^{1-\psi}\right)^{-\frac{1}{1-\psi}}. \tag{A.11}$$

and applying the market clearing condition in the output market:

$$\left(\frac{Y}{\bar{Y}}\right)^{-\frac{1}{\chi}} = c_y = X^{-1}. \tag{A.12}$$

Next, we solve the cost minimization problem within task j:

$$\min_{l(j),k(j)} q(j)k(j) + w(j)l(j) \qquad s.t.\;\; y(j) = \left[(1-\gamma_j)k(j)^{\frac{\nu_j-1}{\nu_j}} + \gamma_j\, l(j)^{\frac{\nu_j-1}{\nu_j}}\right]^{\frac{\nu_j}{\nu_j-1}}. \tag{A.13}$$

Defining the constants $a_i \equiv \gamma_j^{\nu_j}$, $b_j \equiv (1-\gamma_j)^{\nu_j}$, solving (A.13) yields the per-unit cost of input j, $p(j)$:

$$p(j) = \left(a_j w(j)^{1-\nu_j} + b_j q(j)^{1-\nu_j}\right)^{\frac{1}{1-\nu_j}}. \tag{A.14}$$

Plugging (A.10) and (A.14) in the CES Hicksian demand for $k(j)$ and $l(j)$ and imposing labor market clearing gives

$$L(j) = L_0(j)w(j)^{\zeta(j)} = \frac{a_j}{w(j)^{\nu_j}}\left(a_j w(j)^{1-\nu_j} + b_j q(j)^{1-\nu_j}\right)^{\frac{\nu_j-\psi}{1-\nu_j}} X^{\chi-\psi}\bar{Y} = l(j). \tag{A.15}$$

Since $q(j)$ is exogenous, (A.15) provides a condition which implicitly defines $w(j)$. $k(j)$ is then uniquely determined by $w(j)$ and the exogenous $q(j)$:

$$k(j) = \frac{b_j}{q(j)^{\nu_j}}\left(a_j w(j)^{1-\nu_j} + b_j q(j)^{1-\nu_j}\right)^{\frac{\nu_j-\psi}{1-\nu_j}} X^{\chi-\psi}\bar{Y}. \tag{A.16}$$

Taking the ratio of (A.15) and (A.16), we immediately obtain the relative expenditure shares of capital and labor in producing task j:

$$\kappa(j) \equiv \frac{q(j)k(j)}{w(j)l(j)} = \frac{b_j q(j)^{1-\nu_j}}{a_j w(j)^{1-\nu_j}}. \tag{A.17}$$

Implicitly differentiating the above equations thus allows us to perform comparative statics to advances in technology.

A.3 Comparative statics at the task level

We then analyze the impact of technology in producing capital specific to task j on the labor compensation for performing task j. The total effect of innovation on equilibrium wages is the sum of a direct effect on wages holding constant the aggregate productivity X, and an indirect effect which captures the effect of aggregate productivity X on equilibrium wages. All cross-task dependencies operate only through indirect effects.

The direct effect can be obtained by applying the implicit function theorem to (A.15) that

$$-\frac{\partial \log w(j)}{\partial \log q(j)} = \frac{\partial \log w(j)}{\partial \varepsilon(j)} = \frac{(\psi - \nu_j)\kappa(j)}{(\psi + \zeta_j) + (\nu_j + \zeta_j)\kappa(j)} = \frac{\psi - \nu_j}{\nu_j + \zeta_j}\Gamma_j. \tag{A.18}$$

The last equality in (A.18) holds because we define Γ_j as the partial elasticity of the marginal cost of producing task j to changes in the capital price $q(j)$ holding aggregate productivity constant:

$$\Gamma_j \equiv \frac{\partial \log p(j)}{\partial \log q(j)} = \frac{\partial \log p(j)}{\partial \log w(j)}\bigg|_{q(j)} \frac{\partial \log w(j)}{\partial \log q(j)}\bigg|_{X} + \frac{\partial \log p(j)}{\partial \log q(j)}\bigg|_{w(j)} = \frac{(\nu_j + \zeta_j)\kappa(j)}{(\psi + \zeta_j) + (\nu_j + \zeta_j)\kappa(j)}. \tag{A.19}$$

Next, we examine the indirect effect through changes of overall level of industry productivity. Analogous to (A.18), the partial elasticity of the wage level to industry productivity holding $q(j)$ fixed is:

$$A_j \equiv \frac{\partial \log w(j)}{\partial \log X} = \frac{(\chi - \psi)(1 + \kappa_j)}{\psi + \zeta_j + (\nu_j + \zeta_j)\kappa_j}. \tag{A.20}$$

Using (A.18) and (A.20), we can now approximate wage changes due to innovation as

$$\Delta \log w(j) \approx \frac{\partial \log w(j)}{\partial \varepsilon(j)}\varepsilon(j) + \frac{\partial \log w(j)}{\partial \log X}\Delta \log X = \frac{\psi - \nu_j}{\nu_j + \zeta_j}\Gamma_j \varepsilon(j) + A_j \Delta \log X. \tag{A.21}$$

We derive comparative statics for productivity X and other industry aggregate outcomes in section A.7 below.

A.4 Worker earnings growth and technology exposure

Next, we aggregate across tasks to derive predictions for expected changes in incumbent worker earnings in response to innovation. As stated in the main text, we assume that parameters such as Γ_j, A_j do not vary within task type, and denote the corresponding values as Γ_R and A_R for routine tasks and Γ_N and A_N for non-routine tasks.

The level of wage earnings for an individual worker is equal to the total compensation for the tasks she supplies,

$$W(i) = \sum_{j \in J} w(j) l(i,j) I(j \in J_{o(i)}). \tag{A.22}$$

Taking a first order approximation, earnings growth for a worker i is approximately equal to a weighted average of changes in task prices and skills

$$\Delta \log W(i) \approx \sum_{j \in J} s(i,j) \Big(\Delta \log w(j) + \Delta \log l(i,j) \Big), \tag{A.23}$$

where the weights depend on the contribution of each task her occupation performs to her wage

$$s(i,j) = \frac{w(j)l(i,j)I(j \in J_{o(i)})}{W(i)}. \tag{A.24}$$

Note that we assume that workers had similar levels of initial skill in each task they are performing $l(i,j) = \bar{l}(i), \forall j \in J_{o(i)}$. Applying (A.6) and (A.21) we can rewrite (A.23) as

$$\begin{aligned}
\Delta \log W(i) \approx & \sum_{j \in J} s(i,j) \left(\frac{\psi - \nu_j}{\nu_j + \zeta_j} \Gamma_j \varepsilon(j) + A_j \Delta \log X - \left(\beta + \omega \Big[\log \bar{l}(i) - \int \log \bar{l}(i)\, dF(i) \Big] \right) I(j \in J_N) \varepsilon(j) \right) \\
= & \left[\frac{\psi - \nu_R}{\nu_R + \zeta_R} \Gamma_R \right] \xi^R(i) + \left[\frac{\psi - \nu_N}{\nu_N + \zeta_N} \Gamma_N - \beta - \omega \Big[\log \bar{l}(i) - \int \log \bar{l}(i)\, dF(i) \Big] \right] \xi^N(i) \\
& + \Big[(A_R - A_N)\theta(i) + A_N \Big] \Delta \log X
\end{aligned} \tag{A.25}$$

where $F(i)$ denotes the cross-sectional distribution of $\log \bar{l}(i)$ across workers. $\theta(i)$ measures the share of labor compensation to worker i due to performing the routine tasks in her occupation. $\xi^R(i)$ and $\xi^N(i)$ denote the worker's exposure to labor-saving and labor-augmenting technologies, respectively. Also, we define $\tilde{s}^R(i,j)$ and $\tilde{s}^N(i,j)$ as shares normalized to one within each task type:

$$\theta(i) \equiv \sum_{j \in J_R} s(i,j) \tag{A.26}$$

$$\begin{aligned}
\xi^R(i) &\equiv \sum_{j \in J_R} s(i,j)\varepsilon(j) = \theta(i) \sum_{j \in J_R} \frac{s(i,j)}{\sum_{i \in J_R} s(i,j)} \varepsilon(j) = \theta(i) \sum_{j \in J_R} \tilde{s}^R(i,j)\varepsilon(j) \\
\xi^N(i) &\equiv \sum_{j \in J_N} s(i,j)\varepsilon(j) = (1 - \theta(i)) \sum_{i \in J_N} \frac{s(i,j)}{\sum_{i \in J_N} s(i,j)} \varepsilon(j) = (1 - \theta(i)) \sum_{j \in J_N} \tilde{s}^N(i,j)\varepsilon(j).
\end{aligned} \tag{A.27}$$

A.5 Discussion of modeling assumptions

Here, we discuss the sensitivity of our analysis to specific modeling assumptions.

Aggregate skill displacement To simplify the exposition, we have assumed that there is no skill displacement in the aggregate. This assumption has no meaningful impact on our conclusions. In Appendix A.6 we extend the model to allow for skill displacement, which leaves the main implications qualitatively unchanged. The main difference with our baseline model is that, allowing for aggregate skill displacement implies that wages in non-routine tasks are more likely to rise following improvements in labor-augmenting technology due to a labor scarcity effect. This effect is partially offset if effort allocation across tasks is flexible.

Task routineness and capital-labor complementarity Our operating assumption in setting up the model is that technologies that are closely related to an occupation's routine tasks are labor-saving technologies, whereas technologies that are related to an occupation's non-routine tasks likely involve technologies that augment labor. This intuition is in line with most recent work on automation (see, for instance Autor et al., 2003). That said, this modeling assumption mainly affects how we take the model to the data. An alternative interpretation of the model would have been to label the R tasks as low-skill and the N tasks as high-skill, in line with the literature on skill-biased technical change. At a high level, the relevant question is whether technology is labor-saving or labor-augmenting for specific tasks performed by workers. In our empirical section, we also explore an alternative empirical approach that relies on this interpretation.

Other assumptions For simplicity and tractability, we assume that there are frictions to switching occupations which are sufficiently large that we can treat a worker's choice of occupation as if it is fixed in our analysis of comparative statics. Our assumption is in line with the literature (Acemoglu and Restrepo, 2022). Given our empirical focus on relatively short horizons, and the difficulty of observing occupation switching in the data, we view this assumption as reasonable.

A.6 Extension: Adding aggregate skill displacement

In this section, we relax the assumption from the baseline model that there is no aggregate skill displacement. In particular, we assume that the new entrants are not more skilled than incumbents, thus the aggregate labor supply of non-routine task $j \in J_N$ shifts down following a technology innovation, with the same scaling factor β for new entrants and incumbent workers. Therefore, the aggregate labor supply of task j changes following an innovation according to

$$\Delta \log L(j) = \left[\bar{\zeta} + \zeta_j \Delta \log w(j)\right] - \beta I[j \in J_N]\varepsilon(j). \tag{A.28}$$

Compared with (A.7), the new item outside the bracket is due to skill displacement in the aggregate level. The labor supply function can still be expressed as stated in (A.8) and be applied to all equilibrium condition stated in section A.2. The only difference is that the exogenous labor supply shifter $L_0(j)$ now decreases following an innovation such that

$$\frac{\partial L_0(j)}{\varepsilon(j)} = -\beta I[j \in J_N]. \tag{A.29}$$

Correspondingly, (A.18) becomes

$$\frac{\partial \log w(j)}{\partial \log \varepsilon(j)} = \left.\frac{\partial \log w(j)}{\partial \log \varepsilon(j)}\right|_{L_0(j)} + \left.\frac{\partial \log w(j)}{\partial \log L_0(j)}\right|_{q(j)} \frac{\partial \log L_0(j)}{\partial \varepsilon(j)} = \frac{\psi - \nu_j}{\nu_j + \zeta_j}\Gamma_j + \frac{\beta}{\chi - \psi} A_j I[j \in J_N] \tag{A.30}$$

Comparing (A.18) (A.30), we see that all else constant skill displacement will tend to push wages upward following labor augmenting innovations by making labor more scarce relative our baseline model. This labor supply shift partially insulates incumbent workers' wages regardless of the particular choices of the elasticities of substitution ($A_j/(\chi - \psi)$ is strictly positive).

Collecting terms and re-arranging, we have

$$\Delta \log w(j) = \frac{\partial \log w(j)}{\partial \varepsilon(j)}\varepsilon(j) + \frac{\partial \log w(j)}{\partial \log X}\Delta \log X = \left[\frac{\psi - \nu_j}{\nu_j + \zeta_j}\Gamma_j + \frac{\beta}{\chi - \psi} A_j I[j \in J_N]\right]\varepsilon(j) + A_j \Delta \log X \tag{A.31}$$

Analogous to (A.25), we have that

$$
\begin{aligned}
&\Delta \log W(i) \approx \\
&\sum_{j\in J} s(i,j)\left[\left(\frac{\psi-\nu_j}{\nu_j+\zeta_j}\Gamma_j+\frac{\beta}{\chi-\psi}A_j\right)\varepsilon(j)+A_j\Delta\log X-\left(\beta+\omega\left[\log\bar{l}(i)-\int\log\bar{l}(i)\,dF(i)\right]I(j\in J_N)\right)\varepsilon(j)\right] \\
&=\left[\frac{\psi-\nu_R}{\nu_R+\zeta_R}\Gamma_R\right]\xi^R(i)+\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\Gamma_N-\beta(1-\frac{A_N}{\chi-\psi})-\omega\left[\log\bar{l}(i)-\int\log\bar{l}(i)\,dF(i)\right]\right]\xi^N(i) \\
&+\left[(A_R-A_N)\theta(i)+A_N\right]\Delta\log X.
\end{aligned}
\tag{A.32}
$$

Compared to (A.25), the only difference is the $\frac{A_N}{\chi-\psi}$ term, which reflects the impact of labor scarcity on the equilibrium wage.

A.7 Aggregate Effects of Technology Exposure

In this section we examine the changes of aggregate quantiles, including relation between technology exposure and occupation employment, industry productivity, and industry labor share.

A.7.1 Notation at task, occupation, and industry levels

Before discussing comparative statics, we define a number of task- and occupation-level objects which appear in various aggregations below.

We begin at the task level. For task j, the wage bill share $s(j)$, the expenditure share o$s^p(j)$, satisfy

$$
s(j)\equiv\frac{w(j)l(j)}{\sum_{k\in J}w(k)l(k)},\qquad s^p(j)\equiv\frac{p(j)y(j)}{\sum_{k\in J}p(k)y(k)},\qquad \theta\equiv\sum_{j\in J_R}s(j). \tag{A.33}
$$

Given our assumption that $\kappa_j=\kappa_R$ for $j\in J_R$ and $\kappa_j=\kappa_N$ otherwise, we can convert between wage bill share and expenditure shares using our knowledge of the total and task-level labor shares. Specifically,

$$
\frac{s(j)}{s^p(j)}=\frac{w(j)l(j)}{\sum_{k\in J}w(k)l(k)}\frac{\sum_{k\in J}p(k)y(k)}{p(j)y(j)}=\frac{LS(j)}{LS}, \tag{A.34}
$$

where $LS\equiv\frac{\sum_{k\in J}w(k)l(k)}{\sum_{k\in J}p(k)y(k)}$ is the total industry labor share, and $LS(j)\equiv\frac{w(j)l(j)}{w(j)l(j)+k(j)q(j)}=\frac{1}{1+\kappa_j}$ is the task level labor share. Note there are only two possible value for $LS(j)$, thus we define routine and non-routine labor shares via $LS_R\equiv\frac{1}{1+\kappa_R}$ and $LS_N\equiv\frac{1}{1+\kappa_N}$, respectively.

With these definitions in hand, we can properly aggregate across tasks to define analogous occupation-level variables. We define $s(o)$ the aggregate wage bill share of the occupation, and $s_o(j)$ the relative wage bill share of task j within that occupation. The share of the wage bill in routine tasks for a given occupation $\theta(o)$ is defined by

$$
s(o)\equiv\sum_{j\in J_o}s(j),\qquad s_o(j)\equiv\frac{s(j)}{s(o)},\qquad \theta(o)\equiv\sum_{j\in J_o\cap J_R}s_o(j). \tag{A.35}
$$

Next, we denote exposure of labor-saving and labor-augmenting technologies to occupation o $\xi^R(o)$ and $\xi^N(o)$, respectively, in a manner analogous to (A.27) at the worker level:

$$
\xi^R(o)\equiv\sum_{j\in J_o\cap J_R}s_o(j)\varepsilon(j),\qquad \xi^N(o)\equiv\sum_{j\in J_o\cap J_N}s_o(j)\varepsilon(j). \tag{A.36}
$$

$\xi^R(o)$ and $\xi^N(o)$ are weighted average of task-level innovation measures, where the weight on each task is the task-level wage share across workers within the occupation.

Next, we define wage bill and aggregate employment at the occupation level. It is straightforward to define the wage bill for occupation o $WB(o)$, which satisfies

$$WB(o) = \sum_{j \in J_o} w(j)l(j). \tag{A.37}$$

In contrast, because worker productivity is potentially heterogeneous between workers and across different tasks within a given worker, the definition of the number of efficiency units of labor employed in occupation o is not unique. We define the growth rate in the quantity of labor utilized in occupation o using a Divisia index –i.e., as a wage-bill weighted-average of task level employment growth

$$\frac{\Delta L(o)}{L(o)} = \left[\sum_{j \in o} s_o(j) \frac{\Delta l(j)}{l(j)} \right]. \tag{A.38}$$

Finally, we turn to industry-level measures. Further, we define industry-level measures of aggregate exposure of workers to labor-saving and labor-augmenting technologies – $\bar{\xi}^R$ and $\bar{\xi}^N$, respectively – as wage bill-weighted averages of task level technology exposure:

$$\bar{\xi}^R = \sum_{j \in J_R} s(j)\varepsilon(j) = \sum_o s(o)\xi^R(o), \qquad \bar{\xi}^N = \sum_{j \in J_N} s(j)\varepsilon(j) = \sum_o s(o)\xi^N(o) \tag{A.39}$$

Finally, θ is the total wage bill share of all labor-saving tasks, which is defined according to

$$\theta \equiv \sum_{j \in J_R} s(j). \tag{A.40}$$

A.7.2 Comparative statics for occupation-level labor compensation and employment

Using our definitions and comparative statics above, the occupation level wage bill evolves according to

$$\begin{aligned}
\Delta \log WB(o) &\approx \sum_{j \in J_o} s_o(j) \left[\Delta \log w(j) + \Delta \log l(j) \right] \\
&= \sum_{j \in J_o} s_o(j)(1 + \zeta_j)[\frac{\psi - \nu(j)}{\nu(j) + \zeta(j)} \Gamma_j \varepsilon(j) + A_j \Delta \log X] \\
&= \left((\psi - \nu_R) \frac{1 + \zeta_R}{\nu_R + \zeta_R} \Gamma_R \right) \xi^R(o) + \left((\psi - \nu_N) \frac{1 + \zeta_N}{\nu_N + \zeta_N} \Gamma_N \right) \xi^N(o) \\
&+ [A_R(1 + \zeta_R)\theta_o + A_N(1 + \zeta_N)(1 - \theta_o)] \Delta \log X.
\end{aligned}$$

The response of the wage bill across all workers in the same occupation is equal to the change in wages plus the change in employment. Given our definition of employment in (A.38) above, we obtain a closely related comparative static for employment:

$$\begin{aligned}\Delta \log L(o) &\approx \sum_{j\in o} s_o(j)\Delta \log l(j) \\ &= \sum_{j\in o} s_o(j)\zeta(j)\Delta \log w(j) \\ &\approx \zeta_R\left[\frac{\psi-\nu_R}{\nu_R+\zeta_R}\Gamma_R\right]\xi^R(o)+\zeta_N\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\Gamma_N\right]\xi^N(o)+\left[\zeta_R\theta_o A_R+\zeta_N(1-\theta_o)A_N\right]\Delta\log X.\end{aligned} \tag{A.41}$$

The first two terms in equation (A.41) capture the direct effects of labor-saving and labor-augmenting technologies on employment in affected occupations. The last term capture the employment changes due to changes in aggregate industry productivity; we take a closer look at innovation, industry productivity, and other industry outcomes next.

A.7.3 Comparative statics for industry productivity

Here, we explore how industry productivity changes in response to the set of task specific innovation shocks $\varepsilon(j)$ experienced. We apply the implicit function theorem to the second equal sign of the industry market clearing condition (A.12). Specifically, we take the log of both sides of (A.12) and apply the implicit function theorem to take the derivative with respect to $\varepsilon(j)$, which yields the expression

$$\frac{\partial \log X}{\partial \varepsilon(j)} = -\frac{\frac{\partial \log c_y}{\partial \varepsilon(j)}}{1+\frac{\partial \log c_y}{\partial \log X}}. \tag{A.42}$$

Since $\log X = -\log c_y$, we obtain the numerator of (A.42) directly by computing the elasticity of $\log X$ with respect to $\varepsilon(j)$.

$$\frac{\partial \log c_y}{\partial \varepsilon(j)} = -\left(\frac{p(j)}{c_y}\right)^{1-\psi}\frac{\partial \log p_j}{\partial \log q_j} = -s^p(j)\Gamma_j, \tag{A.43}$$

where in (A.43) we apply the fact that

$$\left(\frac{p(j)}{c_y}\right)^{1-\psi} = \frac{p(j)y(j)}{c_y Y} = s^p(j). \tag{A.44}$$

Recall that $s^p(j)$ is the expenditure share of task j.

Next, we implicitly differentiate (A.12), using the definition of aggregate productivity (A.11), with respect to X in order to find the term in the denominator of (A.42),

$$\frac{\partial \log c_y}{\partial \log X} = \sum_{j\in J}\left(\frac{p(j)}{c_y}\right)^{1-\psi}\left[\frac{\partial \log p(j)}{\partial \log w(j)}\frac{\partial \log w(j)}{\partial \log X}\right] = \sum_{j\in J} s^p(j)\frac{A_j}{1+\kappa_j} = LS\left[(A_R-A_N)\theta+A_N\right], \tag{A.45}$$

where ϵ_c, the elasticity of marginal cost of output c_y to an increase in output, satisfies

$$\epsilon_c \equiv \frac{\partial \log c_y}{\partial \log Y} = \frac{1}{\chi}\frac{\partial \log c_y}{\partial \log X} = \frac{LS}{\chi}\left[(A_R-A_N)\theta+A_N\right]. \tag{A.46}$$

Plugging (A.43) and (A.46) into (A.42) implies that, up to a first order approximation, productivity X

changes according to

$$\begin{aligned}
\Delta \log X &\approx \sum_{j\in J} -\frac{\frac{\partial \log c_y}{\partial \varepsilon(j)}}{1+\frac{\partial \log c_y}{\partial \log X}}\varepsilon(j) = \frac{1}{1+\chi\epsilon_c}\sum_{j\in J} s^p(j)\Gamma_j\varepsilon(j) \\
&= \frac{\Gamma_R}{1+\chi\epsilon_c}\frac{LS}{LS_R}\bar{\xi}^R + \frac{\Gamma_N}{1+\chi\epsilon_c}\frac{LS}{LS_N}\bar{\xi}^N .
\end{aligned} \tag{A.47}$$

A.7.4 Comparative statics for industry labor share

In this section, we analyze the evolution of the labor share of output in response to innovation. While the model makes clear predictions about the evolution of the labor share at the task and occupational levels, data are not typically collected at these levels of aggregation. As a result, we focus attention on LS, the aggregate labor share at the industry level.

Differentiating the definition of LS yields the following decomposition:

$$\begin{aligned}
\Delta \log LS &\approx \frac{1}{LS}\sum_{j\in J} s^p(j)LS(j)\left(\Delta \log LS(j) + \Delta \log \frac{p(j)y(j)}{\sum_{k\in J} p(k)y(k)}\right) \\
&= \underbrace{\sum_{j\in J} s(j)\Delta \log LS(j)}_{\text{average within-task LS change}} \underbrace{+\sum_{j\in J} s(j)\left(\Delta \log[p(j)y(j)] - \sum_{k\in J} s^p(k)\Delta \log[p(k)y(k)]\right)}_{\text{between task redistribution}},
\end{aligned} \tag{A.48}$$

where we make use of the identities in (A.34) to simplify the expression. The change in LS is captured by two terms. The first term captures the weighted average change in the labor share within each task, where tasks with higher task shares $s(j)$ get higher weight. The second term captures a reallocation effect, capturing the fact that the aggregate labor share can change if innovation reallocates input expenditures across tasks with different labor share. This term captures a covariance between initial labor shares at the task level and change in shares of input expenditures across different tasks.

First, we consider how the within task term changes with $\varepsilon(j)$, holding industry productivity fixed:

$$\begin{aligned}
\frac{\partial \log LS(j)}{\partial \varepsilon(j)} &= \partial \log\left(\frac{\frac{w(j)l(j)}{q(j)k(j)}}{1+\frac{w(j)l(j)}{q(j)k(j)}}\right)/\partial\varepsilon(j) = [1-LS(j)]\frac{\partial\left(\log[w(j)l(j)] - \log[q(j)k(j)]\right)}{\partial \varepsilon(j)} \\
&= (1-\nu_j)\frac{\psi+\zeta_j}{\nu_j+\zeta_j}\Gamma_j .
\end{aligned} \tag{A.49}$$

where in (A.49) we applied the fact that

$$\frac{\partial(\log[w(j)l(j)] - \log[q(j)k(j)])}{\partial\varepsilon(j)} = (1-\nu_j)\frac{\partial(\log w(j) - \log q(j))}{\partial \varepsilon(j)} = (1-\nu_j)(\frac{\partial \log w(j)}{\partial \varepsilon(j)} + 1) \tag{A.50}$$

by taking log of (A.17) and then taking the derivative with respect to innovation. Following the same approach and differentiating with respect to X, we obtain that

$$\frac{\partial \log LS(j)}{\partial \log X} = (1-\nu_j)[1-LS(j)]A_j . \tag{A.51}$$

Combining (A.49) and (A.51), we obtain an expression how the within-task labor share changes

$$\Delta \log LS(j) \approx \frac{\partial \log LS(j)}{\partial \varepsilon(j)}\varepsilon(j) + \frac{\partial \log LS(j)}{\partial \log X}\Delta \log X = (1-\nu_j)\left[\frac{\psi+\zeta_j}{\nu_j+\zeta_j}\Gamma_j\varepsilon(j) + (1-LS(j))A_j\Delta\log X\right] \tag{A.52}$$

the direction of whom depends on the elasticity of substitution between capital and labor.

Next, we turn to the between-task re-distribution term, which requires us to obtain changes in task expenditure. As before, we take separate derivatives with respect to $\varepsilon(j)$ and X. By taking the log of (A.10) and applying a simple transformation, we obtain that

$$\log p(j)y(j) = (1-\psi)\log p_i + (\chi-\psi)\log X + \log \bar{Y}. \tag{A.53}$$

Differentiating (A.53) with respect to $\varepsilon(j)$, we find that

$$\frac{\partial \log p(j)y(j)}{\partial \varepsilon(j)} = -(1-\psi)\Gamma_i \tag{A.54}$$

and

$$\frac{\partial \log p(j)y(j)}{\partial \log X} = (1-\psi)LS(j)A_j + \chi - \psi. \tag{A.55}$$

Adding up the direct ($\varepsilon(j)$) and indirect ($\Delta \log X$) effects gives us

$$\begin{aligned}\Delta \log p(j)y(j) &\approx \frac{\partial \log p(j)y(j)}{\partial \varepsilon(j)}\varepsilon(j) + \frac{\partial \log p(j)y(j)}{\partial \log X}\Delta \log X \\ &= (1-\psi)\Bigg[-\Gamma_j\varepsilon(j) + A_j LS(j)\Delta\log X\Bigg] + (\chi-\psi)\Delta\log X.\end{aligned} \tag{A.56}$$

where the first term with the brackets depends on elasticity of substitution between tasks. Notice that terms outside the bracket will not contribute to the covariance term since this term is uniform across tasks.

Combining (A.48), (A.52) and (A.56), we can rewrite the change in the total labor share as

$$\begin{aligned}\Delta \log LS \approx \quad & \Gamma_R\left[(1-\nu_R)\frac{\psi+\zeta_R}{\nu_R+\zeta_R} + (\psi-1)\left(1-\frac{LS}{LS_R}\right) + \frac{LS}{LS_R}\frac{\vartheta}{1+\chi\epsilon_c}\right]\bar{\xi}^R \\ &+\Gamma_N\left[\underbrace{(1-\nu_N)\frac{\psi+\zeta_N}{\nu_N+\zeta_N}}_{\substack{\text{within-task substitution}\\ \text{(aggregated direct effect)}}} + \underbrace{(\psi-1)\left(1-\frac{LS}{LS_N}\right)}_{\substack{\text{between-task substitution}\\ \text{(aggregated direct effect)}}} + \underbrace{\frac{LS}{LS_N}\frac{\vartheta}{1+\chi\epsilon_c}}_{\substack{\text{aggregated}\\ \text{indirect effect}}}\right]\bar{\xi}^N,\end{aligned} \tag{A.57}$$

where

$$\begin{aligned}\vartheta = \frac{\partial \log LS}{\partial \log X} = \quad & A_R\theta(1-\nu_R)(1-LS_R) + A_N(1-\theta)(1-\nu_N)(1-LS_N) \\ &+ (1-\psi)\Bigg[A_R\theta(LS_R - LS) + A_N(1-\theta)(LS_N - LS)\Bigg].\end{aligned} \tag{A.58}$$

Examining (A.57), we see that the response of the industry labor share and overall labor-saving $\bar{\xi}_R$ and labor-augmenting $\bar{\xi}_N$ innovation depends on three separate terms. The first two terms inside each one of the brackets multiplying our exposure measures aggregate up the direct effects of productivity improvements on the labor share. The first captures the average change in the labor share (holding X constant) induced, holding the original share of output produced by each task constant. The sign of this first term depends on whether capital and labor are complements or substitutes in performing that task. Given our assumption that $\nu_R > 1$, the within-task reallocation force lowers the labor share for routine tasks, whereas the relationship between ν_N and unity is less obvious.

The second terms within each set of brackets illustrate how technological innovation alters the allocation of expenditures across tasks, holding productivity fixed. The signs of these terms partly depend on the interaction between elasticity of substitution across tasks ψ and the relative labor intensity of routine ($\frac{LS_R}{LS}$) and non-routine ($\frac{LS_N}{LS}$) tasks. The interaction gives that, if tasks are complements ($\psi < 1$), then labor-saving technological improvements—increases in $\bar{\xi}_R$—would shift the share of expenditures towards nonroutine tasks. If it is further the case that routine tasks are more capital-intensive than non-routine tasks, this force would tend to increase the labor share since now high-labor share tasks constitute a larger component of the industry's total expenditures. By contrast, if tasks are sufficient substitutes ($\psi > 1$), then increases in $\bar{\xi}_R$ would induce firms to increase their expenditure towards routine tasks. The labor intensity term also captures how much scope for reallocation there is between two type of tasks.

The last term inside each set of brackets captures indirect effects – the extent to which aggregate productivity changes increase or decrease the labor share. Specifically, ϑ captures the impact of aggregate productivity improvements on the labor share, holding current marginal costs at the task level constant. The sensitivity of the labor share to X ϑ also has within and between effects somewhat analogous to the first two terms in (A.57). The terms outside of the bracket in (A.58) capture within-task substitution between capital and labor and the terms in the bracket capture between-task reallocation associated with X changes, which depends on ψ. If $\psi > 1$, resources get diverted to other tasks and the input share of an unaffected task will shrink. On the contrary, if $\psi < 1$, the input share of an unaffected task will rise, pulling wages up.

A.8 GMM estimation

In this section we outline the GMM estimation of model parameters. We target 16 coefficients to match our closed-form model expressions. We also report the model expressions for all targeted moments together in Table A.7. We provide a detailed discussion of model expressions for targeted moments and their mappings to empirical counterparts, as follows.

A.8.1 Individual responses without industry-level productivity spillover terms

The model-implied worker-level responses without industry-level productivity spillover terms (i.e. when including industry $\times$ year fixed effects) come from equation (A.25). Empirically, we have the technology exposures $\xi^R(i)$ and $\xi^N(i)$ on the right-hand side for the earnings' response regression. The standard deviations for the labor-saving and labor-augmenting technology shock are denoted by $\tilde{\sigma}_R$ and $\tilde{\sigma}_N$ and the CDF of $\log \bar{l}(i)$ is $F(i)$.

$$\begin{aligned}\Delta \log W(i) \approx \quad & \underbrace{\left[\frac{\psi - \nu_R}{\nu_R + \zeta_R}\Gamma_R\right]\tilde{\sigma}_R}_{\text{worker earning to } \xi^R\text{, Homogenous}} \quad \frac{\xi^R(i)}{\tilde{\sigma}_R} \\ & + \left[\underbrace{\left(\frac{\psi - \nu_N}{\nu_N + \zeta_N}\Gamma_N - \beta\right)\tilde{\sigma}_N}_{\text{worker earning to } \xi^N\text{, Homogenous}} \quad \underbrace{-\omega\left[\log \bar{l}(i) - \int [\log \bar{l}(i)]\, dF(i)\right]\tilde{\sigma}_N}_{\text{0 at homogenous level}} \right] \frac{\xi^N(i)}{\tilde{\sigma}_N} + \text{Industry-Time Fixed Effect}\end{aligned} \tag{A.59}$$

Heterogeneous responses of workers' earnings to technology exposure can be derived using the expression above conditional on empirical income sort bins. Since the model implies no heterogeneity in earnings response to labor-saving technologies $\xi^R(i)$ within different income bins, we only examine heterogeneous responses to $\xi^N(i)$, corresponding to the second term in the coefficient on $\xi^N(i)$ in (A.59).

We assume that the CDF of log worker-specific productivity $F(i)$ follows a normal distribution, so that $\log \bar{l}(i) - \int [\log \bar{l}(i)]\, dF(i) \sim N(0,\sigma)$, where we denote the dispersion to be σ; we map the parameter σ to the within-occupation-industry volatility of log earnings in the data. For the heterogeneous coefficients, now consider workers whose earnings are between percentiles a and b in the distribution. The expectation of $\tilde{Z} \sim N(0,\sigma)$ conditional on being in the percentile $[\alpha_a, \alpha_b]$ can be derived as

$$E[\tilde{Z}|\tilde{Z}_a \leq \tilde{Z} \leq \tilde{Z}_b] = \sigma E[Z|Q_a \leq Z \leq Q_b] = \sigma \frac{\int_{Q_a}^{Q_b} z f(z) dz}{b-a} = \sigma \frac{-\int_{Q_a}^{Q_b} f'(z) dz}{b-a} = -\sigma \frac{f(Q_b) - f(Q_a)}{b-a} \quad \text{(A.60)}$$

where Z follows a standard normal distribution, $\tilde{Z}_x$ is the level of $\tilde{Z}$ at x percentile, Φ is the standard normal CDF while f is the standard normal PDF. $Q_x = \Phi^{-1}(x)$. Using (A.60), we have that for the $[a,b]$ income bin, the income response to labor-augmenting technology is

$$\tilde{\sigma}_N \left[\left(\frac{\psi - \nu_N}{\nu_N + \zeta_N} \Gamma_N - \beta \right) + \omega\sigma \frac{f(Q_b) - f(Q_a)}{b-a} \right] \quad \text{(A.61)}$$

We compare the model-implied coefficient on $\xi^R(i)$ and both homogenous and heterogeneous coefficients in responses to technological exposure to their empirical equivalents. For the homogenous coefficient on $\xi^N(i)$ we set $a = 0$ and $b = 1$ in (A.61), which sends the second term to zero and recovers the unconditional coefficient; for the heterogeneous coefficients we examine each of the 5 income bins found in Table 2. We use the model to target the coefficient estimates in column (4), panel B of Table 1 for the homogeneous coefficients; for the heterogeneous coefficients on $\xi^N(i)$ we target the second column of panel B in Table 2. To align with our empirical analysis, we re-scale coefficients to correspond to a shift from the empirical median to 90th percentile of $\xi^R(i)$ or $\xi^N(i)$ by multiplying the standardized coefficients by the difference between the 90th percentiles and medians in the data of the respective standardized empirical measures.

A.8.2 Occupational employment responses

The response of overall occupational employment within the industry to technology exposure can be obtained directly from (A.41). As in the previous section, we also include industry $\times$ year fixed effects in order to net out productivity spillover terms:

$$\Delta \log L(o) \approx \underbrace{\zeta_R \left[\frac{\psi - \nu_R}{\nu_R + \zeta_R} \Gamma_R \right] \tilde{\sigma}_{R_o}}_{\text{employment to } \xi^R(o)} \frac{\xi^R(o)}{\tilde{\sigma}_{R_o}} + \underbrace{\zeta_N \left[\frac{\psi - \nu_N}{\nu_N + \zeta_N} \Gamma_N \right] \tilde{\sigma}_{N_o}}_{\text{employment to } \xi^N(o)} \frac{\xi^N(o)}{\tilde{\sigma}_{N_o}} + \text{Industry-Time Fixed Effect} \quad \text{(A.62)}$$

Here $\tilde{\sigma}_{R_o}$ and $\tilde{\sigma}_{N_o}$ represent the empirical standard deviations for $\xi^R(o)$ and $\xi^N(o)$ within the regression sample for our estimates of (19). We map (A.62) to our estimates in column (2) of Table 3; since we calibrate the model to the 5-year horizon used in our worker-level regressions, we de-annualize the target coefficients from Table 3 by multiplying them by 5. We also again rescale coefficients to correspond to a shift from the median to 90th percentile of exposure within the regression sample to match our empirical scaling of $\xi^R(o)$ and $\xi^N(o)$.

A.8.3 Individual homogenous responses including industry-level productivity spillover terms

In this section we relax the industry $\times$ year fixed effects and examine specifications which instead directly control for industry-level innovative outcomes, as in specification (25). Going forward we assume that specifications involving industry-level aggregates are subject to a random observation error at the industry level–for example, due to potential mismeasurement of overall industry innovation by proxying for it using patent grants. We introduce the following variables: $\bar{\xi}^R_{jt}$ and $\bar{\xi}^N_{jt}$ are the true labor-saving and labor-augmenting technology shocks to industry j at time t, respectively. $\xi^R_{ijt} + \bar{\xi}^R_{jt}$ and $\xi^N_{ijt} + \bar{\xi}^N_{jt}$ are the true labor-saving and labor-augmenting technology shocks to individual j at time t, respectively. Instead of observing $\bar{\xi}^R_{jt}$ and $\bar{\xi}^N_{jt}$, we observe $\bar{\xi}^R_{jt} + e^R_{jt}$ and $\bar{\xi}^N_{jt} + e^N_{jt}$, where e^R_{jt} and e^N_{jt} are random observation errors. The covariance matrix for shock pairs $\begin{pmatrix}\xi^R_{ijt}\\ \xi^N_{ijt}\end{pmatrix}$, $\begin{pmatrix}\bar{\xi}^R_{jt}\\ \bar{\xi}^N_{jt}\end{pmatrix}$ and $\begin{pmatrix}e^R_{jt}\\ e^N_{jt}\end{pmatrix}$ are $\begin{pmatrix}\sigma^2_R & 0\\ 0 & \sigma^2_N\end{pmatrix}$, $\begin{pmatrix}\sigma^2_{\bar{R}} & \sigma_{\bar{R}\bar{N}}\\ \sigma_{\bar{R}\bar{N}} & \sigma^2_{\bar{N}}\end{pmatrix}$, $\begin{pmatrix}\sigma^2_{e^R} & \sigma_{e^R e^N}\\ \sigma_{e^R e^N} & \sigma^2_{e^N}\end{pmatrix}$, respectively. We denote the correlations between the second and third variable pairs $\rho_{\bar{R}\bar{N}}$ and ρ_e, respectively, such that $\sigma_{\bar{R}\bar{N}} = \rho_{\bar{R}\bar{N}}\sigma_{\bar{R}}\sigma_{\bar{N}}$ and $\sigma_{e^R e^N} = \rho_e \sigma_{e^R}\sigma_{e^N}$. All other interactions of random variables have 0 correlation.

We assume that the noise-to-signal ratio is γ for $\bar{\xi}^R_{jt}$ and $\bar{\xi}^N_{jt}$:

$$\sigma^2_{e^S} = \gamma^2 \sigma^2_{\bar{S}}, \qquad S \in \{R, N\} \tag{A.63}$$

We also impose that the industry-level composite $\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt}$ has the same noise-to-signal ratio, which implies[7]

$$\gamma^2 = \frac{Var(e^R_{jt} + e^N_{jt})}{Var(\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt})} = \gamma^2 \frac{\sigma^2_{\bar{R}} + \sigma^2_{\bar{N}} + 2\rho_e \sigma_{\bar{R}}\sigma_{\bar{N}}}{\sigma^2_{\bar{R}} + \sigma^2_{\bar{N}} + 2\rho_{\bar{R}\bar{N}}\sigma_{\bar{R}}\sigma_{\bar{N}}} \tag{A.64}$$

which requires $\rho_e = \rho_{\bar{R}\bar{N}}$, and thus

$$\sigma_{e^R e^N} = \rho_e \sigma_{e^R}\sigma_{e^N} = \gamma^2 \rho_{\bar{R}\bar{N}}\sigma_{\bar{R}}\sigma_{\bar{N}} = \gamma^2 \sigma_{\bar{R}\bar{N}} \tag{A.65}$$

For GMM estimation, we pin down the true volatilities using empirical volatilities and γ. The mapping between empirical volatilities with measurement error and true volatilities is as follows.

$$\begin{aligned}
\tilde{\sigma}^2_{\bar{S}} &\equiv Var(\bar{\xi}^S_{jt} + e^S_j) = (1+\gamma^2)\sigma^2_{\bar{S}}, \qquad S \in \{R, N\}\\
\tilde{\sigma}_{\bar{R}\bar{N}} &\equiv Cov(\bar{\xi}^R_{jt} + e^R_{j,t}, \bar{\xi}^N_{jt} + e^N_{j,t}) = \sigma_{\bar{R}\bar{N}} + \rho_e \sigma_{e^R}\sigma_{e^N} = (1+\gamma^2)\sigma_{\bar{R}\bar{N}}\\
\tilde{\sigma}^2_S &\equiv Var(\xi^S_{ijt} + \bar{\xi}^S_{jt} + e^S_j) = \sigma^2_S + (1+\gamma^2)\sigma^2_{\bar{S}}, \qquad S \in \{R, N\}
\end{aligned} \tag{A.66}$$

Using the expressions in (A.66), we can represent the covariance matrices for $\begin{pmatrix}\xi^R_{ijt}\\ \xi^N_{ijt}\end{pmatrix}$, $\begin{pmatrix}\bar{\xi}^R_{jt}\\ \bar{\xi}^N_{jt}\end{pmatrix}$ and $\begin{pmatrix}e^R_{jt}\\ e^N_{jt}\end{pmatrix}$ using empirical values and γ as

$$\begin{pmatrix}\tilde{\sigma}^2_R - \tilde{\sigma}^2_{\bar{R}} & 0\\ 0 & \tilde{\sigma}^2_N - \tilde{\sigma}^2_{\bar{N}}\end{pmatrix}, \quad \frac{1}{1+\gamma^2}\begin{pmatrix}\tilde{\sigma}^2_{\bar{R}} & \tilde{\sigma}_{\bar{R}\bar{N}}\\ \tilde{\sigma}_{\bar{R}\bar{N}} & \tilde{\sigma}^2_{\bar{N}}\end{pmatrix}, \quad \frac{\gamma^2}{1+\gamma^2}\begin{pmatrix}\tilde{\sigma}^2_{\bar{R}} & \tilde{\sigma}_{\bar{R}\bar{N}}\\ \tilde{\sigma}_{\bar{R}\bar{N}} & \tilde{\sigma}^2_{\bar{N}}\end{pmatrix} \tag{A.67}$$

For the individual regression including productivity spillover terms, the true model implies that

$$y_{ijt} = \beta_1(\xi^R_{ijt} + \bar{\xi}^R_{jt}) + \beta_2(\xi^N_{ijt} + \bar{\xi}^N_{jt}) + \delta(\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt}) + \epsilon_{ijt} \tag{A.68}$$

[7] The model technically implies we should measure the industry-level composite as $\bar{\xi} = \frac{LS}{LS_R}\bar{\xi}^R + \frac{LS}{LS_N}\bar{\xi}^N$; since $\frac{LS}{LS_R}$ and $\frac{LS}{LS_N}$ are both close to one in practice we ignore this nuance and simply take $\bar{\xi} = \bar{\xi}^R + \bar{\xi}^N$.

where

$$\beta_1 = \frac{\psi - \nu_R}{\nu_R + \zeta_R}\Gamma_R, \qquad \beta_2 = \frac{\psi - \nu_N}{\nu_N + \zeta_N}\Gamma_N - \beta, \qquad \delta = \frac{1}{1+\chi\epsilon_c}A_R\Gamma_R \tag{A.69}$$

The last equation imposes $\Gamma_R = \Gamma_N$, $A_R = A_N$, as we do in the model estimation. Responses are normalized by the regressors' empirical volatility.

With the observation error, our empirical estimates yield the potentially biased coefficients $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\delta}$ given by the following regression expression

$$y_{ijt} = \tilde{\beta}_1(\xi^R_{ijt} + \bar{\xi}^R_{jt} + e^R_{jt}) + \tilde{\beta}_2(\xi^N_{ijt} + \bar{\xi}^N_{jt} + e^N_{jt}) + \tilde{\delta}(\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt} + e^R_{jt} + e^N_{jt}) + \tilde{\epsilon}_{ijt} \tag{A.70}$$

As Abel (2019) shows, in the case of multivariate measurement error the coefficient estimates have the following probability limit:

$$\begin{pmatrix} \frac{1}{\tilde{\sigma}_R}\tilde{\beta}_1 \\ \frac{1}{\tilde{\sigma}_N}\tilde{\beta}_2 \\ \frac{1}{\tilde{\sigma}_{\text{comp, with spill}}}\tilde{\delta} \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \delta \end{pmatrix} - (V+\Sigma)^{-1}\Sigma \begin{pmatrix} \beta_1 \\ \beta_2 \\ \delta \end{pmatrix} \tag{A.71}$$

where

$$V = \begin{pmatrix} \sigma^2_R + \sigma^2_{\bar{R}} & \sigma_{\bar{R}\bar{N}} & \sigma^2_{\bar{R}} + \sigma_{\bar{R}\bar{N}} \\ \sigma_{\bar{R}\bar{N}} & \sigma^2_N + \sigma^2_{\bar{N}} & \sigma^2_{\bar{N}} + \sigma_{\bar{R}\bar{N}} \\ \sigma^2_{\bar{R}} + \sigma_{\bar{R}\bar{N}} & \sigma^2_{\bar{N}} + \sigma_{\bar{R}\bar{N}} & \sigma^2_{\bar{R}} + \sigma^2_{\bar{N}} + 2\sigma_{\bar{R}\bar{N}} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma^2_{e^R} & \sigma_{e^Re^N} & \sigma^2_{e^R} + \sigma_{e^Re^N} \\ \sigma_{e^Re^N} & \sigma^2_{e^N} & \sigma^2_{e^N} + \sigma_{e^Re^N} \\ \sigma^2_{e^R} + \sigma_{e^Re^N} & \sigma^2_{e^N} + \sigma_{e^Re^N} & \sigma^2_{e^R} + \sigma^2_{e^N} + 2\sigma_{e^Re^N} \end{pmatrix} \tag{A.72}$$

and $\tilde{\sigma}_{\text{comp, with spill}}$ gives the empirical standard deviation of the industry composite $\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt}$ within the regression sample.

By substituting β_1, β_2 and δ using equation (A.69), and substituting the entries in (A.72) using the empirical mapping shown in (A.67), we can represent the targeted moment $\tilde{\beta}_1$, $\tilde{\beta}_2$ and $\tilde{\delta}$ in terms of the underlying parameters of interest. We only examine the specification with homogenous coefficients, corresponding to the coefficients in the first 3 columns of panel A in Table A.6; finally, to match the scaling of our empirical estimates we further re-scale coefficients β_1 and β_2 so that marginal effects correspond to a shift from the median to 90th percentile of the distributions of exposure in the data.

A.8.4 Industry productivity and labor share responses

For the responses of industry productivity to composite exposure measure, we assume the true expression is

$$y_{jt} = \delta_1(\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt}) + \epsilon_{jt} \tag{A.73}$$

where the model implies

$$\delta_1 = \frac{1}{1+\chi\epsilon_c}\Gamma_R\tilde{\sigma}_{\text{comp}} \tag{A.74}$$

according to (A.47) after imposing $\Gamma_R = \Gamma_N, A_R = A_N$. The scale factor $\tilde{\sigma}_{\text{comp}}$ now gives the standard deviation of the industry-level composite proxy $\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt} + e^R_{jt} + e^N_{jt}$ within this regression sample. As in the previous section, this empirical industry-level aggregates are subject to measurement error, so the empirical specification is instead

$$y_{jt} = \tilde{\delta}_1(\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt} + e^R_{jt} + e^N_{jt}) + \tilde{\epsilon}_{jt} \tag{A.75}$$

With the noise-to-signal ratio γ, the standard expressions for the OLS coefficient under measurement error gives

$$\tilde{\delta}_1 = \frac{Cov(\delta_1(\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt}) + \epsilon_{ij}, \bar{\xi}^R_{jt} + \bar{\xi}^N_{jt} + e^R_{j,t} + e^N_{j,t})}{Var(\bar{\xi}^R_{jt} + \bar{\xi}^N_{jt} + e^R_{j,t} + e^N_{j,t})} = \frac{\delta_1}{1+\gamma^2} \tag{A.76}$$

For the industry labor share response to the industry technology exposure, the true expression is

$$y_{jt} = \eta_1 \bar{\xi}^R_{jt} + \eta_2 \bar{\xi}^N_{jt} + \epsilon_{jt} \tag{A.77}$$

According to (A.57), we have that

$$\begin{aligned} \eta_1 &= \Gamma_R \left[(1-\nu_R)\frac{\psi+\zeta_R}{\nu_R+\zeta_R} + (\psi-1)\left(1-\frac{LS}{LS_R}\right) + \frac{LS}{LS_R}\frac{\vartheta}{1+\chi\epsilon_c} \right] \tilde{\sigma}_{\bar{R}} \\ \eta_2 &= \Gamma_N \left[(1-\nu_N)\frac{\psi+\zeta_N}{\nu_N+\zeta_N} + (\psi-1)\left(1-\frac{LS}{LS_N}\right) + \frac{LS}{LS_N}\frac{\vartheta}{1+\chi\epsilon_c} \right] \tilde{\sigma}_{\bar{N}} \end{aligned} \tag{A.78}$$

Under observation error, we measure $\tilde{\eta}_1$ and $\tilde{\eta}_2$:

$$y_{jt} = \tilde{\eta}_1(\bar{\xi}^R_{jt} + e^R_{j,t}) + \tilde{\eta}_2(\bar{\xi}^N_{jt} + e^N_{j,t}) + \tilde{\epsilon}_{jt} \tag{A.79}$$

Again using the expression from Abel (2019) for coefficient estimates under multivariate measurement error, we have that

$$\begin{pmatrix} \tilde{\eta}_1 \\ \tilde{\eta}_2 \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} - \begin{pmatrix} \sigma^2_{\bar{R}} + \sigma^2_{e^R} & \sigma_{\bar{R}\bar{N}} + \sigma_{e^R e^N} \\ \sigma_{\bar{R}\bar{N}} + \sigma_{e^R e^N} & \sigma^2_{\bar{N}} + \sigma^2_{e^N} \end{pmatrix}^{-1} \begin{pmatrix} \sigma^2_{e^R} & \sigma_{e^R e^N} \\ \sigma_{e^R e^N} & \sigma^2_{e^N} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \frac{1}{1+\gamma^2} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \tag{A.80}$$

which follows immediately from applying the fact that $\rho_e = \rho_{\bar{R}\bar{N}}$ (as shown in equation (A.64)), to get that $\sigma_{\bar{R}\bar{N}} = \gamma^2 \sigma_{e^R e^N}$. Thus the assumption that the noise-to-signal ratios for $\bar{\xi}^R_{jt} + e^R_{j,t}$, $\bar{\xi}^N_{jt} + e^R_{j,t}$, and $\bar{\xi}^R_{jt} + e^R_{j,t} + \bar{\xi}^N_{jt} + e^N_{j,t}$ are all equal to γ implies that the coefficients collapse to the classical univariate measurement error expression here.

We map the model-implied productivity response to the point estimate on composite exposure in column (2) of Table 4; for the labor share responses we compare the model with the two point estimates in column (3) of Table 4. Since all the coefficients from Table 4 are annualized, we multiply the target coefficients from Table 4 by 5 to match the 5-year horizon, like we did for the employment regressions.

The last moment we target is the average labor share LS, which we take from Koh et al. (2020), as well as its standard error. We have that

$$LS = \frac{\sum_{k\in J} w(k)l(k)}{(1+\kappa_R)\sum_{k\in J_R} w(k)l(k) + (1+\kappa_N)\sum_{k\in J_N} w(k)l(k)} = \left[(1+\kappa_R)\theta + (1+\kappa_N)(1-\theta)\right]^{-1} \tag{A.81}$$

A.8.5 Estimation Procedure

We form a parameter vector Θ using the 11 parameters of interest: $\Theta = [\psi, \chi, \nu_R, \nu_N, \zeta_R, \zeta_N, \beta, \kappa_R, \kappa_N, \gamma, \omega]'$. A few other model parameters needed for estimating model moments are pre-calibrated directly from data. The parameter σ controls the standard deviation of in earnings within occupation-industry; this value is 0.531 in the data. We calibrate a routine share of total wage bill (model parameter θ) of 0.340 by taking the total routine wagebill share in the data $\frac{\sum_o \text{Routine Share}_o \times \text{Total Wagebill}_o}{\sum_o \text{Total Wagebill}_o}$ for each of the 1980, 1990, and 2000

IPUMS Samples and 2010 ACS samples, and then averaging across years. The worker-level average routine share is $\theta(i)$ is 0.373, but since we impose $A_R = A_N$, this parameter does not affect our calibration. Finally, the relevant empirical volatilities/covariances are: $\tilde{\sigma}_R = 0.556$, $\tilde{\sigma}_N = 0.822$, $\tilde{\sigma}_{R_o} = 0.582$, $\tilde{\sigma}_{N_o} = 0.804$, $\tilde{\sigma}_{\bar{R}} = 0.435$, $\tilde{\sigma}_{\bar{N}} = 0.650$, $\tilde{\sigma}_{\bar{R}\bar{N}} = 0.229$, $\tilde{\sigma}_{\text{comp, with spill}} = 1.024$, $\tilde{\sigma}_{\text{comp}} = 1.032$.

Let $\hat{F}(\Theta)$ be a function mapping model-implied parameters to the 16 target moments. We choose the parameter vector $\hat{\Theta}$ by minimizing the distance between the model implied $\hat{F}(\Theta)$ and the actual empirical estiamtes F, with the 2 constraints,

$$\begin{aligned} \hat{\Theta} &= \underset{\Theta}{\operatorname{argmin}} \left(F - \hat{F}(\Theta)\right)' W \left(F - \hat{F}(\Theta)\right) \\ s.t. \quad & A_R = A_N \quad \text{and} \quad \Gamma_R = \Gamma_N \end{aligned} \tag{A.82}$$

where the weight matrix W is a diagonal matrix with $W_{ii} = se_i^{-2}$, and se_i is the standard error from estimating moment i.

A.9 Model Application to Artificial Intelligence

A.9.1 Data

Our objective with this exercise is to use model parameters to simulate expected wage growth across occupations in response to cost improvements induced by artificial intelligence technology. We first obtain data on wages and employment from the 2021 American Community Survey and occupation task descriptions from the August 2023 O*NET database update. We then construct proxies for occupational exposure to labor-substituting and labor-augmenting artificial intelligence by following the logic of our model-implied exposure measure as defined in equation (A.27). This requires us to 1) define which tasks could potentially be replaceable by AI (similar to our routine/non-routine designation of DOT tasks for our main analysis) in order to define task boundaries J_R and J_N for automated and augmented tasks, respectively; and, 2) identify the set of tasks for which there currently exist (or likely will soon exist) AI capabilities that are directly applicable to that task as a proxy for tasks j subject to a capital cost improvement of size $\varepsilon_{AI} \equiv -\Delta \log q_{AI}$. To this end, we construct two separate queries to ChatGPT4 designed to separately capture these two concepts.

ChatGPT queries

We use ChatGPT4 March 2023 release to identify if AI is a substitute or complement to occupation tasks:

> A task in category A could be executed by artificial intelligence with only a limited amount of human intervention. A task that is in category B could be completed much more quickly by human labor if they had the assistance of artificial intelligence, but would still require a lot of human intervention and judgment. A task that is in category C would not be substantially affected by artificial intelligence either way. Here is a task description: [insert task description]. Do you think this task is more likely to be in category A, B, or C? In your answer, consider both the current known capabilities of AI technologies as well as their expected capabilities over the next decade. Explain your reasoning in one sentence.

We classify tasks in category A as being substituted by artificial intelligence and the rest as complemented by AI; ONET reports an importance weight of 1 to 5 for each task, which we use to created a weighted average of exposed tasks. Overall, the employment-weighted average share of tasks exposed to substitution by AI is 19.7%. Next, we classify the ease of using AI to assist with those tasks (proxy for task being subject to ε_{AI} shock):

> A task in category A would be highly likely to have low-cost implementations of artificial intelligence-based support for performing this task in the very immediate future and would be likely to be implemented broadly and quickly. A task in category B may potentially have artificial-intelligence supported capabilities, but would currently be more costly and/or less time-saving to implement, and would likely take longer to come into common use. A task in category C would be unlikely to be performed with the assistance of artificial intelligence in the very near future–even if it could be hypothetically possible at some point. Here is a task: [insert task description]. Do you think this task is more likely to be in category A, B, or C? In your answer, consider the present scope of artificial intelligence technology and expected developments over the next decade. Explain your reasoning in one sentence.

We classify tasks in category A as being subject to a capital cost improvement and the remainder as unaffected by AI. Conditional on a task being exposed to automation by AI, the employment-weighted average share of tasks also exposed to a ε_{AI} shock is 82.9%, while the analogous share of tasks classified as exposed to a change improvement of ε_{AI} is 36.7% conditional on the task being complemented by AI.

A.9.2 Constructing Model-Implied Wage Changes Due to AI

Let $s_{i,j}$ denote the compensation share of task j in occupation i (proxied by ONET importance weights); $s^R_{i,j} \equiv \frac{s_{i,j}}{\sum_{i\in J_R} s_{i,j}}$, the compensation share of task j within the set of occupation i's tasks that are substitutable by AI; $s^N_{i,j} \equiv \frac{s_{i,j}}{\sum_{i\in J_N} s_{i,j}}$, the compensation share of task j within the set of occupation i's tasks that are not substitutable by AI; $\theta(i)_{AI} = \sum_{j\in J_R} s_{i,j}$ the share of occupation i tasks substitutable by AI; and $I(j)$, an indicator of whether task j is subject to a technological improvement due to AI. We then calculate the exposure of occupation i to being substituted and complemented, respectively, by AI as follows:

$$\xi^R_{AI}(i) = \theta(i)_{AI} \times \sum_{j\in J_R} s^R_{i,j} \times I(j) \times \varepsilon_{AI} \tag{A.83}$$

$$\xi^N_{AI}(i) = (1-\theta(i)_{AI}) \times \sum_{j\in J_N} s^N_{i,j} \times I(j) \times \varepsilon_{AI} \tag{A.84}$$

We then use the wage equation (A.25) and GMM estimates of model parameters—subtracting off the aggregate spillover term $[(A_R - A_N)\theta(i) + A_N]\Delta \log X$—to get predicted wage growth for each occupation relative to any aggregate productivity improvements. We calibrate the degree of cost improvement ε_{AI} using the episode of order clerks around the turn of the 21st century examined in our prior work (Kogan, Papanikolaou, Schmidt, and Seegmiller, 2021). Order clerks were particularly hard hit by the rise of e-commerce, with many of their tasks being automated away by information technologies that made electronic purchasing easier; consequently, their wages declined by about 20 percent relative to other clerking occupations over a period of roughly 10 years. We set this event as the upper bound on the impact of AI (which is likely quite conservative). Matching this 20% decline to a 5-year horizon, we accordingly calibrate ε_{AI} so that the most exposed occupation experiences a 10% decline in wages due to technological exposure to AI (before any skill displacement effects are taken into account) over the next 5 years.

B Data Appendix

B.1 Converting Patent and Occupation Task Texts for Numerical Analysis

Here, we briefly overview our conversion of unstructured patent text data into a numerical format suitable for statistical analysis. We obtain text data for measuring patent/job task similarity from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. We use the patent text data parsed from the USPTO patent search website in Kelly et al. (2021), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET.[8] We then combine all tasks for a given occupation at the 2010 O*NET 6-digit level into one occupation-level corpus. The process for cleaning and preparing the text files for numerical representation follows the steps outlined below.

We first clean out all non-alphabetic characters from the patent and task text, including removing all punctuation and numerical characters. We then convert all text to lowercase. At this stage each patent and occupation-level task text are represented by a single string of words separated by spaces. To convert each patent/occupation into a list of associated words we apply a word tokenizer that separates the text into lists of word tokens which are identified by whitespace in between alphabetic characters. Since many commonly used words carry little semantic information, we filter the set of tokens by first removing all "stop words"– which include prepositions, pronouns, and other common words carrying little content–from the union of several frequently used stop words lists.

Stop words come from the following sources:

- https://pypi.python.org/pypi/stop-words
- https://dev.mysql.com/doc/refman/5.1/en/fulltext-stopwords.html
- http://www.lextek.com/manuals/onix/stopwords1.html
- http://www.lextek.com/manuals/onix/stopwords2.html
- https://msdn.microsoft.com/zh-cn/library/bb164590
- http://www.ranks.nl/stopwords
- http://www.text-analytics101.com/2014/10/all-about-stop-words-for-text-mining.html
- http://www.webconfs.com/stop-words.php
- http://www.nltk.org/book/ch02.html (NLTK stop words list)

We also add to the list of stop words the following terms that are ubiquitous in the patent text but don't provide information regarding the content and purpose of the patent: abstract, claim, claims, claimed, claiming, present, invention, united, states, patent, description, and background. The final stop word list contains 1337 unique terms that are filtered out.

[8]The DOT to SOC crosswalk is available at https://www.onetcenter.org/crosswalks.html. Since the time we originally obtained the crosswalk, O*NET has subsequently replaced the 2010 SOC code–based version we use in this paper with a crosswalk derived from the 2019 SOC code scheme.

Even after removing stop words, we expect much of the remaining text to offer little information regarding the purpose and use of a given patent or the core job functions expected to be performed by workers in a given occupation. In order to focus on the parts of the document most likely to contain relevant information, we retain descriptive and action words–i.e. nouns and verbs–and remove all other tokens. We do this using the part-of-speech tagger from the NLTK Python library. Finally, we lemmatize all remaining nouns and verbs, which is to convert them to a common root form. This converts all nouns to their singular form and verbs to their present tense. We use the NLTK WordNet Lemmatizer to accomplish this task. After these steps are completed, we have a set of cleaned lists of tokens for each patent and each occupation's tasks that we can then use to compute pairwise similarity scores.

B.2 Classifying tasks based on their capital/labor complementarity

Our model emphasizes the distinction between exposure to labor-saving technologies (technologies that substitute for labor in performing certain tasks) and labor-augmenting technologies (technologies that augment labor in performing certain tasks). A key issue then is to how empirically identify these two distinct sets of tasks. Conceptually, tasks in which capital can substitute for labor are tasks that can be performed by capital (automation). To that end, we rely on two complementary ideas in the literature. The first is the notion of routine tasks (Acemoglu and Autor, 2011); the key idea here is that routine tasks are the tasks that can be *potentially* be performed by machines. By contrast, non-routine tasks cannot be performed by machines; instead workers these machines to increase their productivity when performing these tasks. Second, we rely on the notion of skill-biased technical change, in which there are a certain type of tasks that require low levels of related skill that can be performed by machines, while capital is a complement to labor when performing high-skill tasks (Krusell et al., 2000; Goldin and Katz, 2008).

B.2.1 Classification into routine and non-routine tasks

To implement this distinction in the data we need to first partition the set of tasks for each occupation into routine and non-routine tasks. Our text data on job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. Given that the main empirical analysis in this paper focuses on worker-level outcomes in the post-1980 sample, we use the task descriptions from the 1991 DOT—which is mostly identical to the 1977 DOT version beyond the addition of some IT-related occupations—rather than more recent versions from O*NET. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we use a crosswalk from DOT occupations to the considerably coarser and yet still detailed set of 6-digit Standard Occupation Classification (SOC) codes from O*NET. We then combine all tasks for a given occupation at the 2010 SOC 6-digit level into one occupation-level corpus.

We then use GPT4 to classify individual tasks (defined by distinct sentences in DOT task descriptions) into routine or non-routine:

> A routine task can be defined as follows: A routine task involves carrying out a limited and well-defined set of work activities, those that can be accomplished by following explicit rules. These tasks require methodical repetition of an unwavering procedure, and they can be exhaustively specified with programmed instructions and performed by machines. Tell me whether the following task is primarily routine, primarily non-routine, or involves a mix of both routine and non-routine tasks; and, explain your reasoning in one sentence. [insert sentence describing task].

We apply this query for each one of the sentences of each occupation's task description in the DOT. Out of a total of approximately 90,000 sentences, GPT4 characterizes 62% of these as referring to routine tasks, 15% as non-routine and 22% as a mixture of routine and non-routine. We group the latter two categories into a non-routine category.

We validate the output produced by GPT4 with a routine task intensity (RTI) index, which we construct from the six standardized occupational task types for the 2010 ONET using code provided by Acemoglu and Autor (2011). Specifically, we take the sum of the standardized routine manual and routine cognitive scores and subtract off the standardized non-routine manual (interpersonal), non-routine manual (physical), non-routine cogntive (analytical), and non-routine cognitive (interpersonal) scores generated by the code from Acemoglu and Autor (2011).[9] To compare the two, we calculate the average share of tasks at the occupation level that are classified as routine by GPT4. Panel A of Figure A.1 plots the RTI index versus the average share of routine tasks at the occupation level. We see that the two are highly correlated (81%) and the relation is approximately linear.

At this point, it is worth emphasizing that we view the 'routineness' of a task as an inherent property of a task that is invariant to the current state of technology. In this regard, we should think of routine tasks as those that can be potentially automated. The rate of arrival of breakthrough innovations determines which tasks are actually automated. Thus, for example, booking airline tickets is a routine task, but it required advances in communication before it could be (mostly) automated. That said, it is possible that the answers provided by GPT4 are influenced by the current state of technology.

B.2.2 Classification into high- and low-skill tasks

As a robustness check, we also explore an alternative classification of tasks to build our technology exposure measures that relies on the idea that low-skill tasks are more easily automated than high-skill tasks (Krusell et al., 2000). Given that the definition of skill can be somewhat vague, we use a definition that relies on related experience provided by ONET, termed Specific Vocational Preparation (SVP),

> Specific Vocational Preparation is the amount of lapsed time required by a typical worker to learn the techniques, acquire the information, and develop the facility needed for average performance in a specific job-worker situation. Tell me whether attaining proficiency in the below occupation task requires A) an extensive amount (more than 5 years); B) a fair amount (1 to 5 years); C) a moderate amount (3 months to 1 year); or D) very little (less than 3 months) of specific vocational preparation; and, explain your reasoning in one sentence. [insert sentence describing task].

As before, we apply this query for each one of the sentences in an occupation's description of tasks. ChatGPT classifies 61% of these task sentences in the D category, with the remainder classified as C(25%), B (13%), and A (0.01%). We label tasks in category D as low experience while tasks in categories A/B/C as high experience.

We validate our alternative measure using ONET's job zone classification, which assigns occupations a score between 1 (lowest) and 5 (highest) depending on the extent of their required preparation, which includes education requirements, related experience, and on-the-job training. As we see in Panel B of Figure A.1, there is a strong correlation (85%) between the share of tasks classified as high-experience by GPT4 and ONET's job classification.

We conclude that our classification of occupation tasks into routine/non-routine or high/low skill captures economically meaningful variation that can be used to construct measures of technology exposure in line with our model. The correlation between these two sets of classification is relatively high, but not perfect: out of the tasks that are classified as routine, approximately 80 percent are classified as low experience. By contrast,

[9] This is similar to what Autor et al. (2003) do to create a routine-task intensity index using the prior Dictionary of Occupational Titles occupation task scores. They take the log of routine score and subtract off the log of abstract and log of manual scores; we do not take logs since there are zeroes in the raw ONET scores, and we also follow Acemoglu and Autor (2011) in standardizing and centering the measures around zero.

out of the tasks that are classified as non-routine, approximately 70 percent are classified as high required experience. We primarily rely on the routine/non-routine classification as our baseline measure, and relegate the results using the high/low skill classification to the appendix.

B.3 Measuring the distance between patents and worker tasks using word embedding vectors

We identify technologies that are relevant to specific worker groups as those that are similar to the descriptions of the tasks performed by a given occupation. We do so by analyzing the textual similarity between the description of the innovation in the patent document and the worker's job description.

To identify the similarity between a breakthrough innovation and an occupation, we need to identify meaningful connections between two sets of documents that account for differences in the language used. The most common approach for computing document similarity is to create a matrix representation of each document, with columns representing document counts for each term (or some weighting of term counts) in the dictionary of all terms contained in the set of documents, and with rows representing each document. Similarity scores could then be computed simply as the cosine similarity between each vector of weighted or unweighted term counts:

$$Sim_{i,j} = \frac{V_i}{||V_i||} \cdot \frac{V_j}{||V_j||} \tag{B.1}$$

Here V_i and V_j denote the vector of potentially weighted terms counts for documents i and j.

This approach is often referred to as the 'the bag-of-words' approach, and has been used successfully in many settings. For example, Kelly et al. (2021) use a variant of this approach to construct measures of patent novelty and impact based on pairwise distance measures between patent documents. Since patent documents have a structure and a legalistic vocabulary that is reasonably uniform, this approach works quite well for patent-by-patent comparisons. However, this approach is less suited for comparing patent documents to occupation task descriptions. These two sets of documents come from different sources and often use different vocabulary. If we were to use the bag-of-words approach, the resulting vectors V_i and V_j would be highly sparse with most elements equal to zero, which would bias the distance measure (B.1) to zero.

The root cause of the problem is that the distance measure in (B.1) has no way of accounting for words with similar meanings. For example, consider a set of two documents, with the first document containing the words 'dog' and 'cat' and the other containing the words 'puppy' and 'kitten'. Even though the two documents carry essentially the same meaning, the bag of words approach will conclude that they are distinct: the representation of the two documents is $V_1 = [1, 1, 0, 0]$ and $V_2 = [0, 0, 1, 1]$, which implies that the two documents are orthogonal, $\rho_{1,2} = 0$.

To overcome this challenge, we leverage recent advances in natural language processing that allow for synonyms. The main idea behind this approach is to represent each word as a dense vector. The distance between two word vectors is then related to the likelihood these words capture a similar meaning. In our approach, we use the word vectors provided by Pennington, Socher, and Manning (2014), which contains a vocabulary of 1.9 million word meanings (embeddings) represented as (300-dimensional) vectors. The two most popular approaches are the "word2vec" method of Mikolov, Sutskever, Chen, Corrado, and Dean (2013) and the global vectors for word representation introduced by Pennington et al. (2014). These papers construct mappings from extremely sparse and high-dimensional word co-occurence counts to dense and comparatively low-dimensional vector representations of word meanings called word embeddings. Their word vectors are highly successful at capturing synonyms and word analogies (vec(king) − vec(queen) ≈ vec(man) − vec(woman) or vec(Lisbon) − vec(Portugal) ≈ vec(Madrid) − vec(Spain), for example). Thus they are well-suited for numerical representations of the "distance" between words. The word vectors provided by Pennington et al. (2014) are trained on 42 billion word tokens of web data from Common Crawl and are available at https://nlp.stanford.edu/projects/glove/.

To appreciate how our metric differs from the standard bag-of words approach it is useful to briefly examine how word embeddings are computed in Pennington et al. (2014). Denote the matrix X as a $V \times V$ matrix of word co-occurence counts obtained over a set of training documents, where V is the number of words in the vocabulary. Then $X_{i,j}$ tabulates the number of times word j appears in the context of the word i.[10] Denote $X_i = \sum_k X_{i,k}$ as the number of times any word appears in the context of word i, and the probability of word j occuring in the context of word i is $P_{i,j} \equiv X_{i,j}/X_i$. The goal of the word embedding approach is to construct a mapping $F(\cdot)$ from some d-dimensional vectors x_i, x_j, and $\tilde{x}_k$ such that

$$F(x_i, x_j, \tilde{x}_k) = \frac{P_{i,k}}{P_{j,k}} \tag{B.2}$$

Imposing some conditions on the mapping $F(\cdot)$, they show that a natural choice for modeling $P_{i,k}$ in (B.2) is

$$x_i^T \tilde{x}_k = \log(X_{i,k}) - \log(X_i) \tag{B.3}$$

Since the mapping should be symmetric for i and k they add "bias terms" (essentially i and k fixed effects) which gives

$$x_i^T \tilde{x}_k + b_i + b_k = \log(X_{i,k}) \tag{B.4}$$

Summing over squared errors for all pairwise combinations of terms yields the weighted least squares objective

$$\text{Min}_{x_i, \tilde{x}_k, b_i, b_k} \sum_{i=1}^{V} \sum_{j=1}^{V} f(X_{i,j}) \left(x_i^T \tilde{x}_k + b_i + b_k - \log(X_{i,j}) \right)^2 \tag{B.5}$$

Here the observation-specific weighting function $f(X_{i,j})$ equals zero for $X_{i,j} = 0$ so that the log is well defined, and is constructed to avoid overweighing rare occurrences or extremely frequent occurrences. The objective (B.5) is a highly-overidentified least squares minimization problem. Since the solution is not unique, the model is trained by randomly instantiating x_i and $\tilde{x}_k$ and performing gradient descent for a pre-specified number of iterations, yielding d-dimensional vector representations of a given word. Here d is a hyper-parameter; Pennington et al. (2014) find that $d = 300$ works well on word analogy tasks.

Since (B.5) is symmetric it yields two vectors for word i, x_i and $\tilde{x}_i$, so the final word vector is taken as the average of the two. The ultimate output is a dense 300-dimensional vector for each word i that has been estimated from co-occurence probabilities and occupies a position in a word vector space such that the pairwise distances between words (i.e. using a metric like the cosine similarity) are related to the probability that the words occur within the context of one another and within the context of other similar words. Note that the basis for this word vector space is arbitrary and has no meaning; distances between word embeddings are only well-defined in relation to one another and a different training instance of the same data would yield different word vectors but very similar pairwise distances between word vectors.

The next step consists of using these word vectors to construct measures of document similarity. To begin, we first construct a weighted average of the word embeddings with a document (a patent or occupation description). Specifically, we represent each document as a (dense) vector X_i, constructed as a weighted

[10] Pennington et al. (2014) use a symmetric 10 word window to determine "context" and weight down occurences that occur further away from the word (one word away receives weight 1, two words away receives weight 1/2, etc.).

average of the set of word vectors $x_k \in A_i$ contained in the document,

$$X_i = \sum_{x_k \in A_i} w_{i,k} x_k. \tag{B.6}$$

A key part of the procedure consists of choosing appropriate weights $w_{i,k}$ in order to emphasize important words in the document. We also note that there is not a Pennington et al. (2014) word embedding estimate for every word in every patent. For example, occasionally there are OCR text recognition errors in patent documents, which can lead to misspelled words that do not have an embedding estimate as a result; there may also be patents that use extremely specific and technical terms (i.e. chemical patents introducing an new and very specific molecule) which do not have a word embedding available. In such cases we simply calculate (B.6) for the set of words which do have a Pennington et al. (2014) embedding estimate—which is the vast majority of terms in practice.

In natural language processing, a common approach to emphasize terms that are most diagnostic of a document's topical content is the 'term-frequency-inverse-document-frequency' (TF-IDF). In brief, $TFIDF_{i,k}$ overweighs word vectors for terms that occur relatively frequently within a given document and underweighs terms that occur commonly across all documents. We follow the same approach: in constructing (B.6), we weigh each word vector by

$$w_{i,k} \equiv TF_{i,k} \times IDF_k. \tag{B.7}$$

The first component of the weight, term frequency (TF), is defined as

$$TF_{i,k} = \frac{c_{i,k}}{\sum_j c_{i,j}}, \tag{B.8}$$

where $c_{i,k}$ denotes the count of the k-th word in document i—a measure of its relative importance within the document. The inverse-document frequency is

$$IDF_k = \log\left(\frac{\text{\# of documents in sample}}{\text{\# of documents that include term } k}\right). \tag{B.9}$$

Thus, IDF_k measures the informativeness of term k by under-weighting common words that appear in many documents, as these are less diagnostic of the content of any individual document. We compute the inverse-document-frequency for the set of patents and occupation tasks separately, so that patent document vectors underweight word embeddings for terms appearing in many patents and occupation vectors underweight word embeddings for job task terms that appear in the task descriptions of many other occupations.

Armed with a vector representation of the document that accounts for synonyms, we next use the cosine similarity to measure the similarity between patent i and occupation j:

$$\text{Sim}_{i,j} = \frac{X_i}{||X_i||} \cdot \frac{X_j}{||X_j||} \tag{B.10}$$

This is the same distance metric as the bag of words approach, except now X_i and X_j are dense vectors carrying a geometric interpretation akin to a weighted average of the semantic meaning of all nouns and verbs in the respective documents.

To illustrate the difference between our approach and the standard bag of words, consider the following example of two documents, with the first document containing the words 'dog' and 'cat' and the other containing the words 'puppy' and 'kitten'. Even though the two documents carry essentially the same meaning, the bag of words approach will conclude that they are distinct: the representation of the two documents is

$$V_1 = [1, 1, 0, 0], \quad \text{and} \quad V_2 = [0, 0, 1, 1] \tag{B.11}$$

which implies that the two documents are orthogonal, $\rho_{1,2} = 0$. Here, the TF-IDF weights in our simple example satisfy $TF_{1,dog} = 1/2$ and $IDF_{dog} = \log(2)$, with similar logic applying to "cat"; this proceeds analogously for document 2 containing "puppy" and "kitten".

By contrast, in the word embeddings approach, these two documents are now represented as

$$X_1 = (1/2) \times \log(2) x_{dog} + (1/2) \times \log(2) x_{cat} \tag{B.12}$$

and similarly for X_2. Here x_{dog}, x_{cat} would have been trained using the Pennington et al. (2014) method described above on a very large outside set of documents. Hence, in this case since word vectors are estimated such that $x_{dog} \approx x_{puppy}$ and $x_{cat} \approx x_{kitten}$, we now have $\text{Sim}_{1,2} \approx 0.81$ using the word vectors estimated by Pennington et al. (2014). A weighted average word embedding approach has been shown in the natural language processing literature to achieve good performance on standard benchmark tests for evaluating document similarity metrics relative to alternative methods that are much more costly to compute (see, e.g. Arora, Liang, and Ma, 2017). A relative disadvantage is that it ignores word ordering—which also applies to the more standard 'bag of words' approach for representing documents as vectors. However, since we have dropped all stop words and words that are not either a noun or a verb, retaining word ordering in our setting is far less relevant.

In sum, we use a combination of word embeddings and TF-IDF weights in constructing a distance metric between a patent document (which includes the abstract, claims, and the detailed description of the patented invention) and the detailed description of the tasks performed by occupations. Our methodology is conceptually related, though distinct, to the method proposed by Webb (2020), who also analyzes the similarity between a patent and O*NET job tasks. Webb (2020) focuses on similarity in verb-object pairs in the title and the abstract of patents with verb-object pairs in the job task descriptions and restricts his attention to patents identified as being related to robots, AI, or software. He uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings. By contrast, we infer document similarity by using geometric representations of word meanings (GloVe) that have been estimated directly from word co-occurence counts. Furthermore, we use not only the abstract but the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention. In addition to employing a different methodology, we also have a broader focus: we are interested in constructing time-series indices of technology exposures. As such, we compute occupation-patent distance measures for all occupations and the entire set of USPTO patents since 1836.

Last, we perform several adjustments to the raw measure of similarity (14). First, we remove yearly fixed effects. We do so in order to account for language and structural differences in patent documents over time.[11] Second, we impose sparsity: after removing the fixed effects we set all patent $\times$ occupation pairs to zero that are below the 80th percentile in this fixed-effect adjusted similarity. This imposes that the vast majority of patent–occupation pairs are considered unrelated to one another, and only similarity scores sufficiently high in the distribution receive any weight. Last, we scale the remaining non-zero pairs such that a patent/occupation pair at the 80th percentile of yearly adjusted similarities has a score equal to zero and the maximum adjusted score equals one. We denote by $\rho_{i,j}$ the adjusted similarity metric between patent j

[11] Patents have become much longer and use much more technical language over the sample period, and the OCR text recognition of very early patents is far from perfect.

and occupation i.

While we compute $\rho_{i,j}$ for all patent–occupation pairs, we restrict to the set of patents identified as breakthroughs by the KPST procedure when we construct our time-varying measures occupation-industry exposure to technology—as in the main text equation (15). Additionally, though we initially calculate $\rho_{i,j}$ for DOT documents with occupations defined at the 2010 SOC code level, we compute η and ξ using the modified Census occ1990 scheme (commonly referred to as "occ1990dd" codes) from Autor and Dorn (2013). The occ1990dd codes are comparable over time and can be readily linked to the CPS or Census, which proves useful for our analysis of occupational- and worker-level outcomes across various datasets. After crosswalking SOC occupations to the occ1990dd level, we take the average of the original 2010 SOC version of $\rho_{i,j}$ within each linked occ1990dd occupation code; we then use the resulting occ1990dd-level measure of $\rho_{i,j}$ in equation (15).

B.4 Validation

We validate our technology exposure measures using a large language model (GPT4). In particular, our operating assumption thus far has been that, technology that relates to routine tasks is likely a substitute for workers (labor-saving) while technology that relates to workers' non-routine tasks is likely to be a complement (labor-augmenting), a view consistent with Autor et al. (2003) and Acemoglu and Autor (2011). However, it is not obvious whether our methodology for capturing distance between innovations and tasks can accurately identify these two types of technology. The examples in the previous section do suggest that it is possible to isolate between labor-saving and labor-augmenting technologies, but they do not constitute a full validation.

To validate these assumptions we use GPT4. Given the high cost of performing this query, we focus on a random sub-sample of 10,000 breakthrough patents issued from 1980 to 2007, evenly distributed across years, and restrict the text input to the abstract of the patent and the occupation's DOT title.

First, we examine whether GPT4 agrees that a technology's distance to routine tasks indicates that it is a substitute. In particular, for each breakthrough patent in the sub-sample, we focus on the five occupations most closely related to that patent in terms of the routine task content of their job description, as well as the five occupations least related—using the raw similarity metric in equation (14). We then query GPT4, for each one of these 100,000 patent-occupation pairs, whether the technology is likely to be a substitute for workers' tasks,

> Here is the abstract of a patent: [Patent abstract here]. Here are some tasks: [DOT title here]. Do you think the technology mentioned above can perform some tasks mentioned above formerly performed by workers? Output yes or no, and your reasoning in one sentence.

We then examine whether GPT4 agrees that distance to non-routine tasks indicates complementarity. Using the same 10,000 breakthrough patents, we now isolate the five most- and the five least-related occupations to each patent in terms of their non-routine tasks. For each of these patent-occupation pairs, we query GPT4 whether the technology in question is likely to complement workers in that occupation

> Here is the abstract of a patent: [Patent abstract here]. Here are some tasks: [DOT title here]. Do you think the patent mentioned above can increase the productivity of workers when performing some of the tasks mentioned above? Output yes or no, and your reasoning in one sentence.

Overall, we see in columns (1) and (3) of Table A.1 that GPT4 largely agrees with our classification. Specifically, for the five most closely related occupations in terms of their routine tasks, GPT4 agrees that this technology is likely to be labor-saving in 86% of the cases, while it only agrees with 4% of the cases for the least exposed occupations. By contrast, when focusing on the distance to non-routine tasks, GPT4 agrees in 83% of the cases that this technology is likely to be complementary to workers in the five most closely related occupations, while only in 10% of the cases for the least exposed occupations. Columns (2) and (4) report the results of the same validation exercise for our high/low required experience measure, which are largely comparable.

B.5 BLS Industry Productivity Data

We obtain data on 4-digit NAICS industry outcomes from the Bureau of labor Statistics (found here: `https://www.bls.gov/productivity/tables/labor-productivity-detailed-industries.xlsx`). The data cover 1987-2022, though we are constrained to the 1987-2012 period based off availability of our industry innovation measures. BLS total factor productivity estimates are only available for manufacturing industries, so we take the industry per-unit labor costs index as our analogue of industry productivity in the model (this corresponds to the "Unit Labor Costs" series from the BLS data). This also has the benefit of matching well with our model-based notion of industry productivity (A.11), which is the per-unit input cost of final industry output. The cost-based productivity series is an index normalized to 100 within industry in the year 2012, and growth in costs are computed relative to this within-industry baseline. We also take changes in the log levels of the BLS labor share and labor compensation series to measure growth in industry labor shares and wagebills, respectively. Finally, since we control for log employment and weight by employment shares in the regressions, we use the BLS "hours worked" series to measure industry employment.

B.6 Publicly Available IPUMS Census Survey Data

While we make use of restricted access data sources (described further in the next subsection) for our primary analysis in sections 3 and 4 of the main text, we also construct a version of the measure that can be created using publicly available sources, which we utilize to summarize variation in our measure in Figure 1 and Figure 2. Specifically, we use the 5% samples from the 1980, 1990, and 2000 Censuses obtained from the Integrated Public Use Microdata Series (IPUMS) database.[12] We construct wages and individual labor supply weights following code from Acemoglu and Autor (2011). For this data we cannot use our direct links of patents to NAICS industries, and the best time series consistent industry code is the Census ind1990 designation. Hence we use probabilistic weights from 3-digit cooperative patent technology classes to NAICS industries created by Goldschlag et al. (2016). We obtain the extended version which includes service industries and can be found here. Goldschlag et al. (2016) provide links at the 2007 NAICS level, so we exploit the fact that IPUMS provides both 2007 NAICS designations (variable "indnaics") and time-series consistent ind1990 codes in the 2008-2012 ACS, which we use to create a correspondence between ind1990 and NAICS codes; we then use this correspondence to create probabilistic assignments of patents to ind1990 codes for each of the Census survey years; finally, weighting patents by industry assignment probabilities allows us to construct a version of our main measure (15) that varies at the occ1990dd-by-ind1990-by-year level. We then merge this onto our IPUMS sample and compute employment-weighted averages of exposure at the occupation-by-year level to generate Figure 1 and Figure 2, as described in the main text.

[12] While in principle we could supplement with ACS survey data for later years, our measure cuts off in 2007 and so we use the 1980, 1990, and 2000 Censuses to preserve a consistent 10-year sample interval.

B.7 Census administrative data

B.7.1 DER-CPS Sample

We use a random sample of individual workers tracked by the Current Population Survey (CPS) and their associated Detailed Earnings Records (DER) from the Census—which contains their W2 tax income. We limit the sample to individuals who are older than 25 and younger than 55 years old. Our sample includes individuals coming from the 1981-1991, 1994, and 1996-2016 ASEC waves for whom valid individual identifiers (a Census Protected Individual Key) can be assigned. In selecting which years to include in our sample, we exactly follow the labor force attachment restrictions imposed by Braxton, Herkenhoff, Rothbaum, and Schmidt (2021).

The CPS includes information on demographic information such as age and gender, but more importantly occupation at the time of the interview. We assign workers to occupations based on their response to the CPS survey (CPS "occ" variable). We construct a crosswalk between the yearly CPS occupations codes and the occ1990dd classification scheme and assign all CPS occupations their corresponding occ1990dd code. We assign this occupation to the worker for the next 3 years, thus effectively dropping observations where the CPS interview date is older than 3 years—so that the occupation information is relatively recent. Workers' employers are differentiated by the federal Employer Identification Number (EIN); if workers report earnings from more than one EIN in a calendar year then we assign them the EIN of highest W2 earnings in that year. Our final worker-year earnings panel spans the years 1981-2016.

B.7.2 Assigning Industry of Origination to Patents

We merge the individual worker records from the CPS-DER matched sample to patent data at the industry (4-digit NAICS) level. Specifically, we identify the industry of patent origination by relying on the Census SSEL patent–assignee database, which provides a corresponding SSEL firm identifier ("firmid") for each patent, which we then use to obtain the firm's 4-digit NAICS code. In particular, we use two SSEL patent–assignee crosswalks: the newer Business Dynamics Statistics of Patenting Firms database (BDS-PF) and an older patent-SSL crosswalk created by Kerr and Fu (2008). The BDS-PF links are available starting with the 2000 SSEL. We use the BDS-PF firmid-patent links for any patents for which it is available. Otherwise we use the union of links created by Kerr and Fu (2008) from the 1976–1999 SSL data. We obtain NAICS codes by using the "firmid" identifier to join with the Longitudinal Business Database (LBD), and we use the 2012 version of Fort and Klimek (2018) NAICS codes, which are adjusted to improve industry comparability over time. In cases where a firmid matches to multiple NAICS codes we apply the 4-digit NAICS code of highest employment based off the LBD. Finally, we drop from our analysis patents that are assigned to the 2-digit NAICS code 55, "Management of Companies and Enterprises", since this code consists of firms that manage and hold controlling interests in other companies, making the NAICS code uninformative about the industry where the actual production is taking place.

B.7.3 Decennial and ACS Surveys

For the analysis in Section 4.1, we use restricted access versions of Census Decennial surveys in 1980, 1990, and 2000; because ACS surveys have much smaller sample size, we average across ACS years 2008-2012 for observations in the year 2010. These surveys have the advantage of using the actual earnings responses rather than top-coded earnings.

In order to construct a crosswalk between the Census industry codes available in each of these survey years in the restricted-access survey data, we first obtain the mapping between the time-consistent Census

1990 industry scheme found in publicly available IPUMS surveys and the 2012 NAICS coding scheme. We use the 2013-2017 5-year ACS to do this because it is the first 5-year survey period containing 2012 NAICS codes. Because the 1990 Census industry scheme is available for all years, we can then construct a mapping between Census industry codes and 4-digit NAICS codes, which is the level of industry that we use throughout the main analysis in the paper. Because some NAICS codes are aggregated in the mapping to the 1990 Census industry scheme, we have to consolidate both Census and NAICS codes to create consistent mapping that is as close as possible to the original 4-digit NAICS codes. After consolidating we are left with 188 modified 4-digit NAICS codes from the original 288.

Following the age restrictions in our worker-level analysis, we include individuals between the ages of 25 to 55 in our sample. We weight employment by the product of survey sample weights and individuals' survey-reported hours worked per week, and weeks worked per year. If hours or weeks worked are reported in intervals for a given survey year, we take the median of that interval. As in our CPS-based analysis, we crosswalk occupations to the time-consistent occ1990dd coding scheme throughout. We then take occupation-industry aggregates of each outcome and control in each survey year. Specifically, we take the total labor employment weights, and the average wage, age, female population share, and years of education, all weighted by employment.

B.7.4 Longitudinal Business Database

We merge the DER-CPS sample with the Longitudinal Business Database (hereafter LBD, Jarmin and Miranda (2002)) to get worker industry information and data on employers' wages and revenues. We assign workers to a consistently-defined industry code based on their employer's industry codes in the LBD. To do this we use a deduplicated crosswalk from LBD firm identifiers ("firmid") to firm EINs, and we assign each EIN with the Fort and Klimek (2018) 2012 4-digit NAICS code based off LBD industry information. We compute firm-level wages by aggregating yearly LBD payroll and employment information to the EIN level, and we compute firm average wages for a given year by taking the ratio of payroll to total employment. Since LBD annual payroll information is reported for the last 12 months as of March of that year, while W2 earnings are reported over the calendar year, we use the average of the LBD-implied wage for years t and year $t + 1$ as a measure of employers' average wages in year t. We use the 2016 of edition of the LBD to compute firm average wages and to obtain firm industry information. Finally, we utilize the newer revenue-enhanced LBD (Haltiwanger, Jarmin, Kulick, Miranda, and Penciakova, 2019) to measure revenue productivity, defined as revenues per worker; for similar reasons as for firm wages, we take the average of revenues per worker in years t and $t + 1$ to proxy for revenue productivity in year t. We have LBD information on firm industries and wages from the start of our sample in 1981 through 2015, while revenue information is available from 1997 through 2016.

B.7.5 Restrictions for Assigning Worker Earnings Bins

To allow the effects to vary with prior income, we assign workers into five groups based on their income in the previous year compared to workers in the same occupation and NAICS4 industry. These groups are defined based on the following percentiles of prior income [0%, 25%), [25%, 50%), [50%, 75%), [75%, 95%), [95%, 100%] calculated within industry–occupation cells. In the (uncommon) case when NAICS4 industry-occupation cells have fewer than 10 individuals, we broaden the industry definition from 4-digit NAICS to 2-digit NAICS. Any Industry–Occupation cells that still have fewer than 10 individuals after moving to 2-digit NAICS codes are dropped.

B.8 Details on Worker Earnings Variance Decompositions and Analysis of Labor Revenue Productivity

In this appendix section we provide details on calculations discussed in section 3 of the main text; these include a variance decomposition of worker earnings into between- and within-occupation-industry/firm type components and correlating firm revenue productivity with worker earnings rank.

B.8.1 Earnings Decompositions

We perform two variance decompositions of log W2 earnings using our baseline DER-CPS sample, one which includes just occupation-industry components and another that adds on a firm component. Our first wage decomposition is as follows:

$$\overline{\log\left(\text{W2 Earnings}_{i,t}\right)} = \alpha_{occ(i,t)\times ind(i,t),t} + \epsilon_{i,t} \tag{B.13}$$

Here i indexes an individual worker and $\alpha_{occ(i,t)\times ind(i,t),t}$ denotes yearly fixed effects for worker i's occupation-industry pair at time t, and $\overline{\log\left(\text{W2 Earnings}_{i,t}\right)}$ denotes cross-sectionally demeaned log earnings. Similar to the restrictions we impose for assigning earnings rank bins, we use a worker's occ1990dd $\times$ 4-digit NAICS code for $\alpha_{occ(i,t)\times ind(i,t),t}$ if there are 10 or more observations in the yearly occ-ind cell, and 2-digit NAICS code if fewer than 10 observations. We drop an occ-ind cell if there are still fewer than 10 observations after moving to 2-digit NAICS codes. Estimating (B.13) results in the following variance decomposition of annual log earnings:

$$\text{Var}_t\left(\overline{\log\left(\text{W2 Earnings}_{i,t}\right)}\right) = \underbrace{\text{Var}_t\left(\hat{\alpha}_{occ(i,t)\times ind(i,t),t}\right)}_{\text{Between occupation-industry}} + \underbrace{\text{Var}_t\left(\hat{\epsilon}_{i,t}\right)}_{\text{Within occupation-industry}} \tag{B.14}$$

We plot the time-series of this variance decomposition in panel A of A.2. The time series starts in 1981 when our CPS-DER earnings panel begins, and ends in 2015 when the LBD industry information ends. The variance of both the between- and within–occupation-industry components increases over time; over the whole time period the within occupation-industry component explains 57.7% of the variance in log earnings, and this ratio never falls below 55% for any year in the sample period.

Next we examine an extension on the decomposition in (B.13) by adding dummies for firm type:

$$\overline{\log\left(\text{W2 Earnings}_{i,t}\right)} = \alpha_{occ(i,t)\times ind(i,t),t} + \alpha_{\text{FirmRank}(i,t),t} + \epsilon_{i,t} \tag{B.15}$$

Here $\alpha_{\text{FirmRank}(i,t),t}$ are yearly dummies for 100 wage bins based off annually ranking workers on their employers' LBD average wages; these dummies capture wage effects from being employed at a high-paying firm. The earnings variance decomposition now becomes

$$\begin{aligned}\text{Var}_t\left(\overline{\log\left(\text{W2 Earnings}_{i,t}\right)}\right) = {} & \underbrace{\text{Var}_t\left(\hat{\alpha}_{occ(i,t)\times ind(i,t),t}\right)}_{\text{Between occupation-industry}} + \underbrace{\text{Var}_t\left(\hat{\alpha}_{\text{FirmRank}(i,t),t}\right)}_{\text{Between firm-type}} + \\ & \underbrace{2\times\text{Cov}_t\left(\hat{\alpha}_{occ(i,t)\times ind(i,t),t}, \hat{\alpha}_{\text{FirmRank}(i,t),t}\right)}_{\text{Sorting by occ-ind and firm-type}} + \underbrace{\text{Var}_t\left(\hat{\epsilon}_{i,t}\right)}_{\text{Within occupation-industry and firm-type}}\end{aligned} \tag{B.16}$$

In panel B of A.2 we graph the time series of the earnings decomposition in (B.16). The sample for this panel ends in 2014 instead of 2015 since we average the LBD-implied firm average wage over two adjacent years (as explained in appendix B.7), and our LBD firm wage coverage cuts off in 2015. The residual component $\mathrm{Var}_t(\hat{\epsilon}_{i,t})$ now represents the variance of earnings within both occupation-industry and firm type; the variance of this within component again increases throughout the sample period. The full period average share of variance of the within occupation-industry and firm type component is now 48.4%; this implies that accounting for firm type effects and sorting can only explain about 16.1% (i.e. 1 - 48.4/57.7) of the within occupation-industry variance share of log earnings.

We also report a similar variance decomposition for college attainment by estimating a version of (B.13), where a dummy for being a college grad replaces worker earnings as the dependent variable. We find that 56.1% of variation in college attainment can be explained within occupation-industry cells over the full sample period.

B.8.2 Revenue Productivity of Employers and Worker Earnings Ranks

Next we describe in detail our calculations from section 3 on how worker earnings rank within occupation-industry correlates with firm labor productivity. In particular, we estimate the following:

$$\log\left(\text{Revenue/Worker}_{i,t}\right) = \alpha_{EarningsRank(i,t)} + \alpha_{occ(i,t)\times ind(i,t),t} + \epsilon_{i,t} \tag{B.17}$$

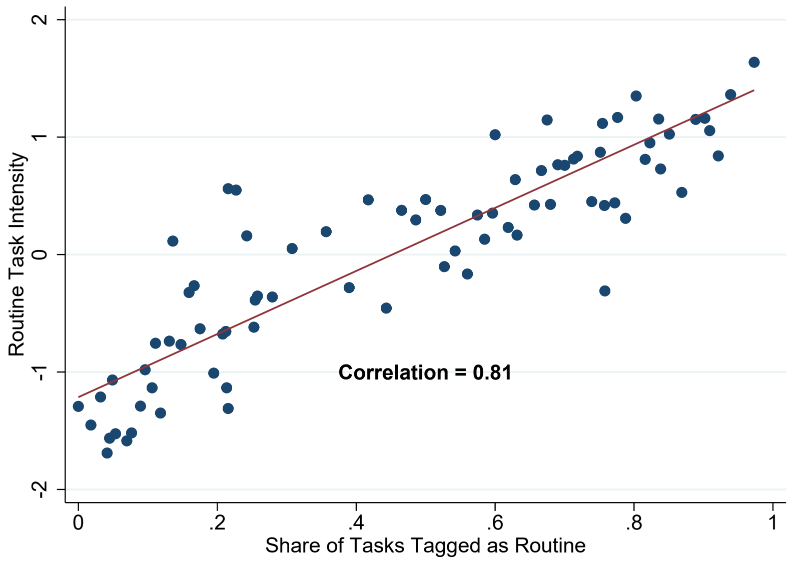
We use the revenue-enhanced version of the LBD to calculate $\log\left(\text{Revenue/Worker}_{i,t}\right)$, the log revenues per employee of individual i's year-t employer. The term $\alpha_{EarningsRank(i,t)}$ denotes dummies for worker i earnings rank at time t. These correspond to the same earnings bins (i.e. [0%, 25%), [25%, 50%), [50%, 75%), [75%, 95%), [95%, 100%]) as throughout the paper. We also control for yearly occupation-industry fixed effects so that comparisons are within occupation and industry. We use the [0%, 25%) bin as the omitted category. While our main sample spans 1981-2016, the sample period for this regression is constrained by availability of LBD revenues data begins in 1997. The estimates among the top four income bins are as follows: 0.122 (SE = 0.012); 0.231 (SE=0.019); 0.345 (SE=0.026); and, 0.454 (SE=0.034) for the [25%, 50%), [50%, 75%), [75%, 95%), and [95%, 100%] income rank bins, respectively.[13] Hence we find that workers in the highest occupation-industry income bin tend to be employed at firms with about 0.45 higher log revenue productivity than workers in the bottom bin, and this estimate is highly statistically significant. Looking at the distribution of revenue productivity in Table A.2, this corresponds to a difference of about 0.4 standard deviations. The revenue productivity effects are also increasing in worker earnings rank across all income bins, suggesting a strong monotonic relationship across the earnings distribution.

[13] We report standard errors clustered by industry.

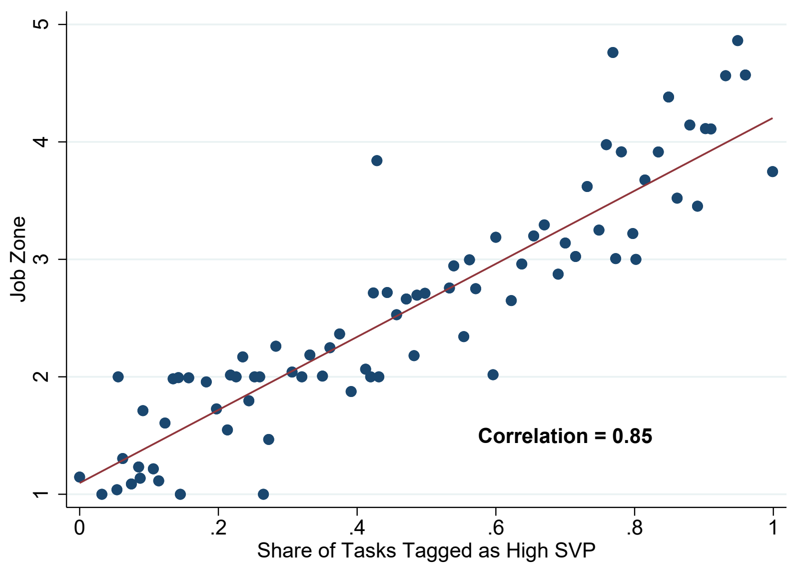
Appendix Figures and Tables

Figure A.1: Validation of ChatGPT4 split task classification

A. Validation of task split into Routine/Non-routine using RTI



B. Validation of task split into low/high related experience using ONET job zones



Note: This figure compares occupation routine task intensity scores constructed from Acemoglu and Autor (2011) task categories, with the share of occupation tasks from the Dictionary of Occupational Titles (DOT) tagged as routine by ChatGPT in panel A; panel B compares the share of tasks tagged as high experience require with ONET job zone category, which varies between 1 and 5, with 5 corresponding to high amounts of prior preparation and/or experience required. We weight 6-digit SOC occupations by employment shares from Acemoglu and Autor (2011). See appendix B.2 for further details on the classification of tasks into routine/non-routine or low/high experience and the construction of the data in these figures.

Figure A.2: Dispersion in worker earnings: decomposition

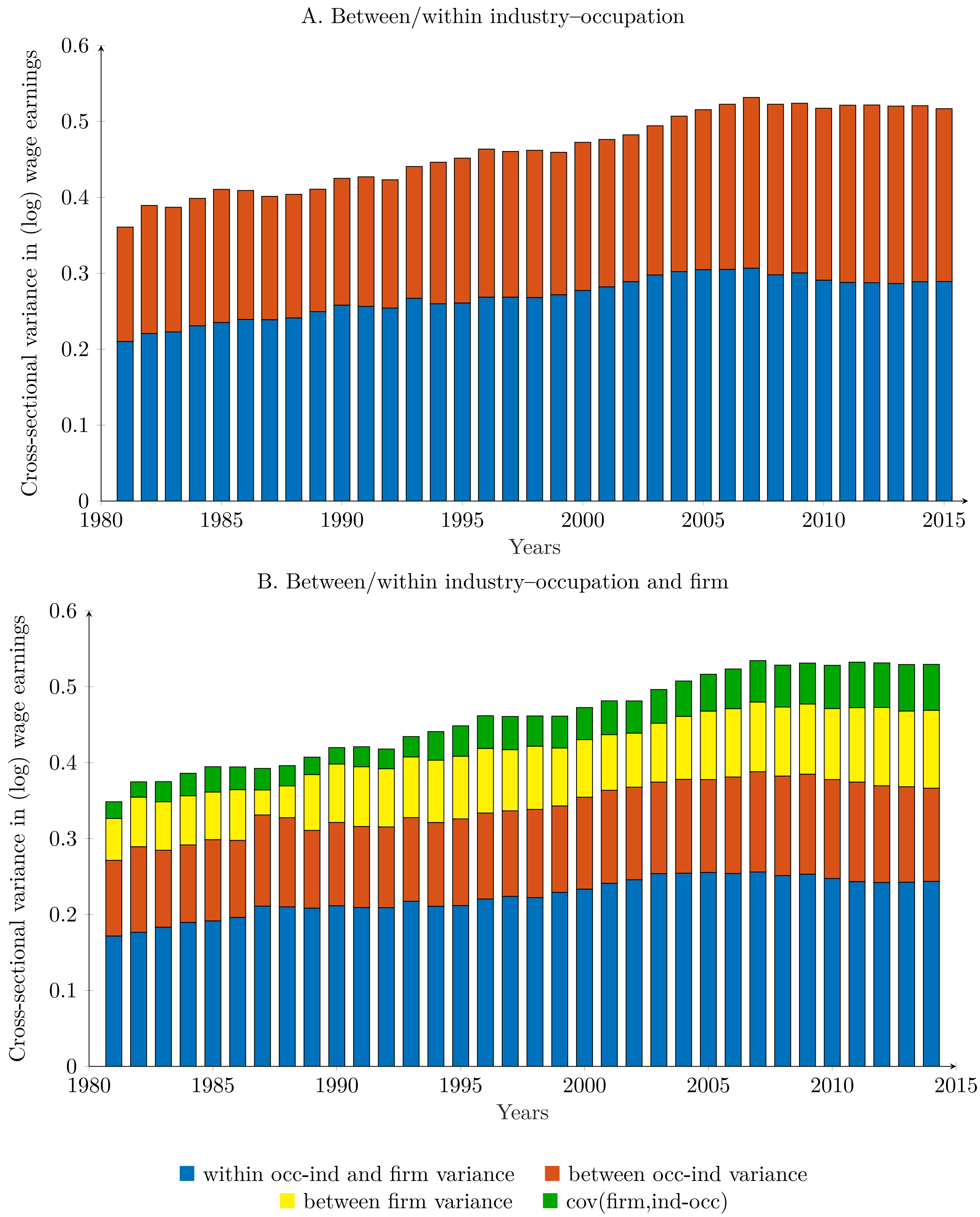
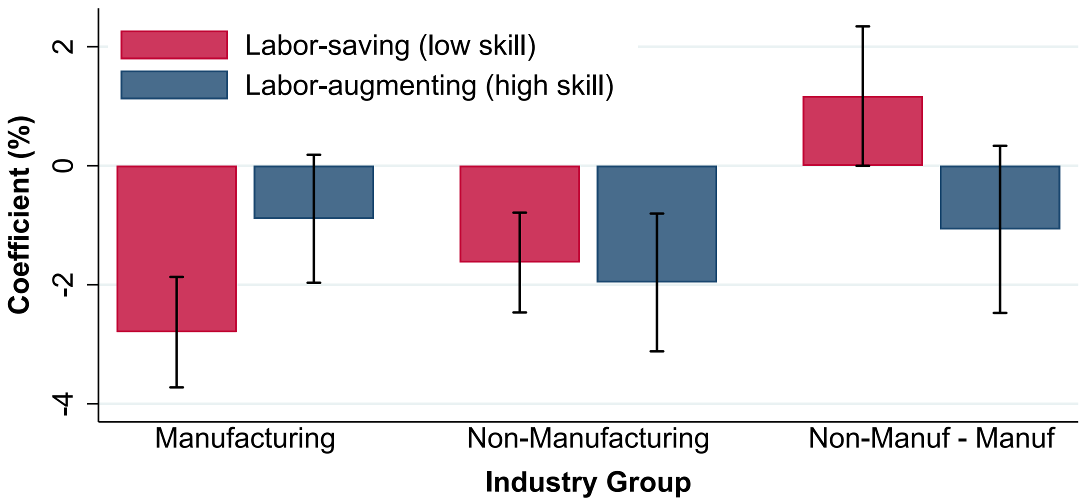


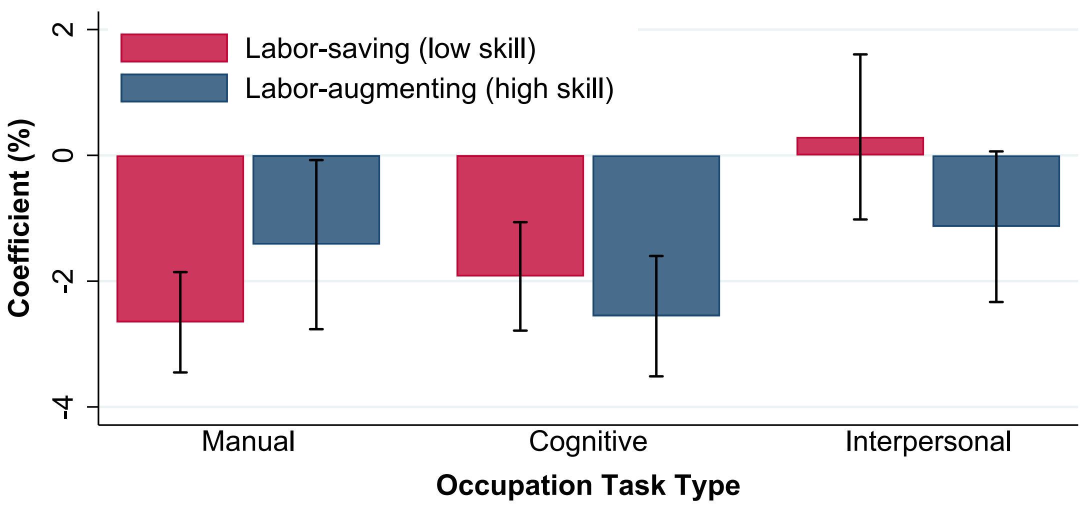
Figure graphs the share of variance in log earnings into within- and between- occupation-industry components (panel A); and within- and between- occupation-industry and firm-type components (panel B) for each year in our sample. See Appendix B.8 for details on these decompositions.

Figure A.3: Technology exposure and worker earnings growth, by industry or occupation type (alternative measure based on low/high skill tasks)

A. Manufacturing vs Non-Manufacturing

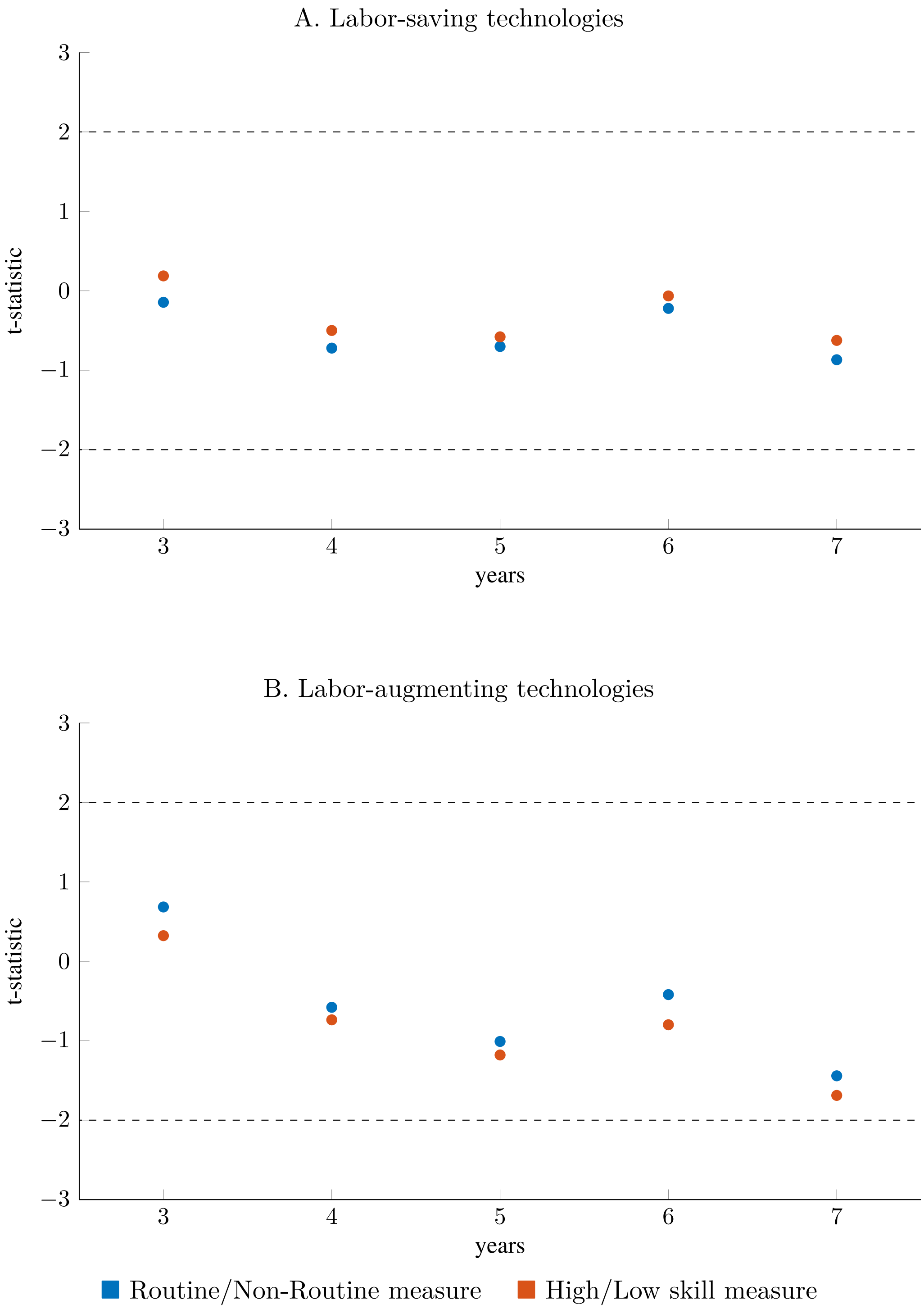


B. Occupation Task Type: Manual/Cognitive/Interpersonal



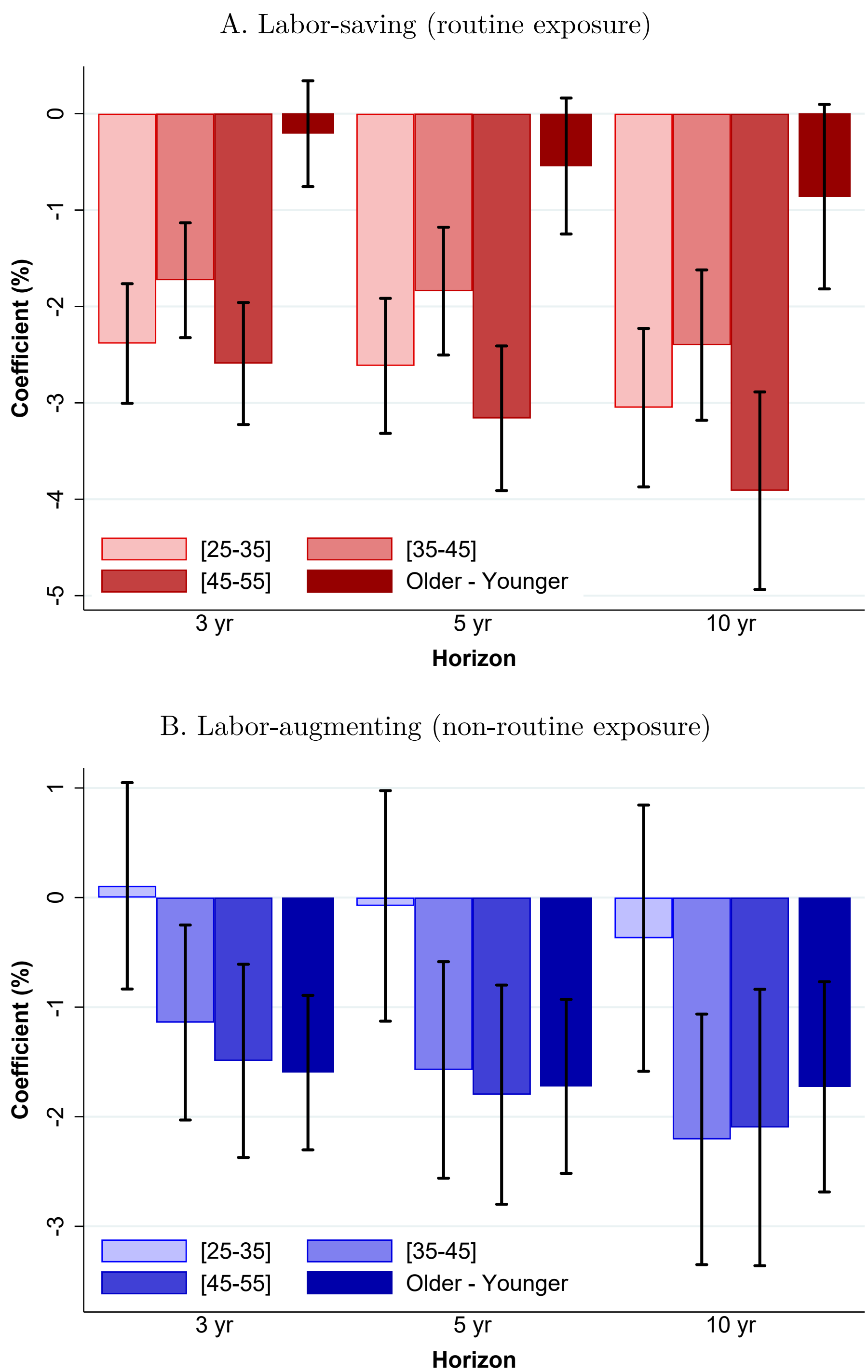
Note: Figure shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text, except now we use ChatGPT to classify individual tasks as requiring either low or high specific preparation required; we allow these coefficients to vary by industry type (panel A) or occupation task type (panel B). In panel A we compare coefficients for individuals employed in or out of manufacturing (broadly defined as 2-digit NAICS codes 11 through 33); in panel B we designate occupations as primarily focusing on either manual, cognitive, or interpersonal tasks using task scores from Acemoglu and Autor (2011). The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We plot 95% confidence intervals from standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. All specifications include industry × year, occupation × year, within occupation–industry income bin × year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Figure A.4: Wage growth and future technology exposure



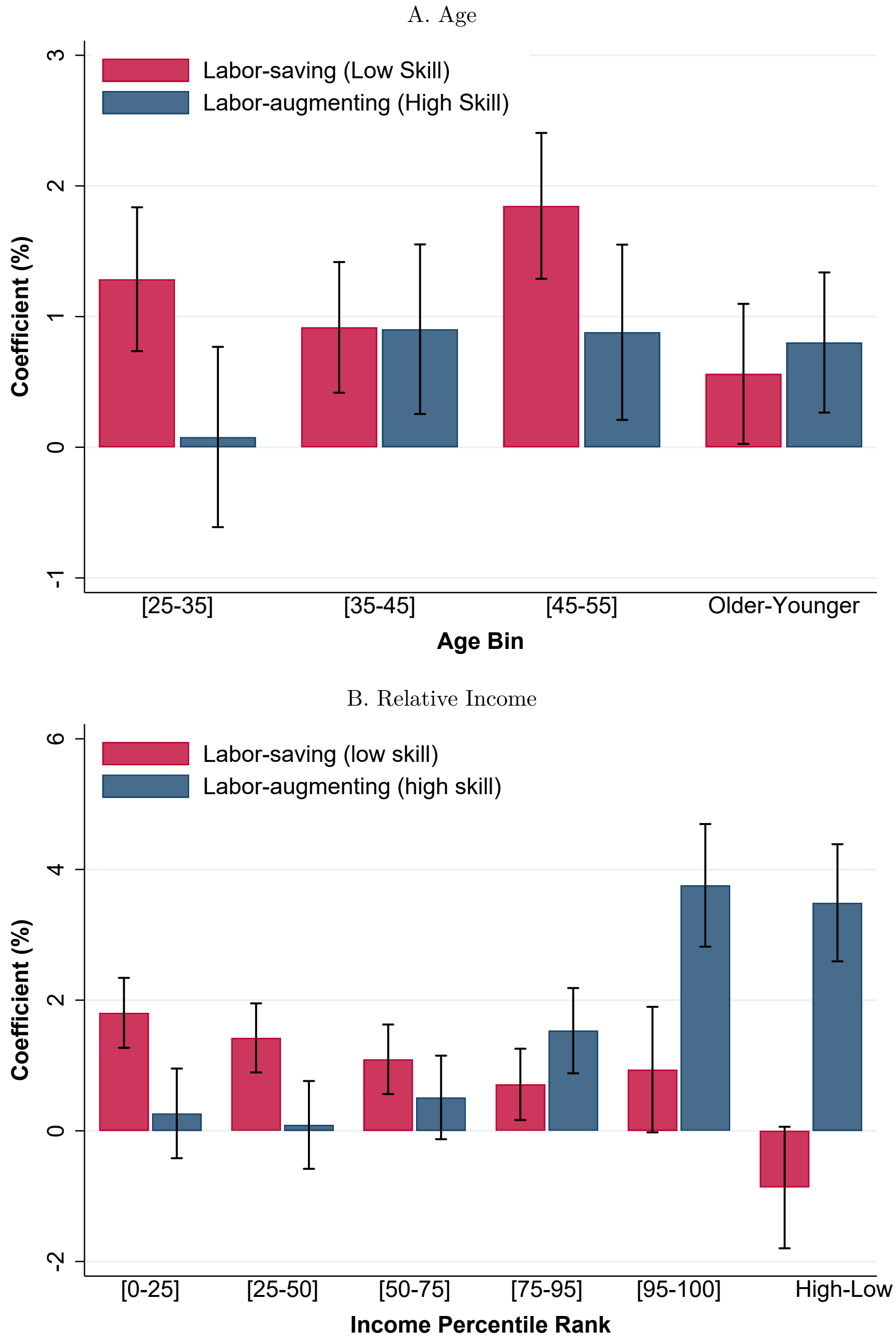
Note: Plot reports the t-statistic corresponding to the partial correlation of wage growth Δw^i_{t+5} and future technology exposure for horizons $h = 1 \ldots 7$. Specifically, we orthogonalize both variables with respect to the right-hand side variables of equation (16) using our most saturated specification and report the t statistic of a regression of the orthogonalized component of our future innovation measure on past wage growth.

Figure A.5: Technology exposure and worker earnings growth, by worker age and horizon



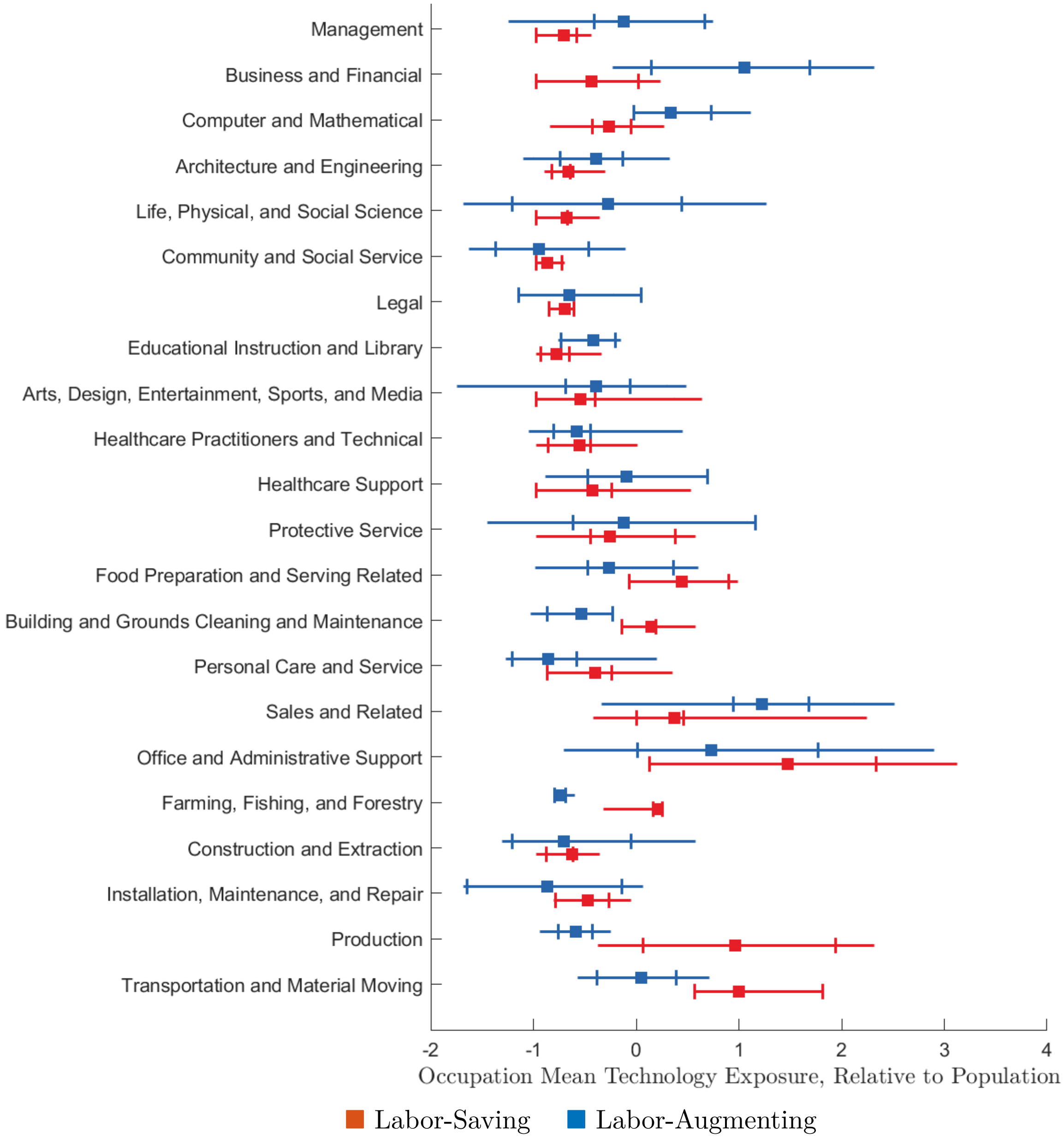
Note: Figure shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text, where we allow these coefficients to vary by age group and for 3-, 5-, and 10-year horizons. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the given time horizon. We plot 95% confidence intervals from standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. All specifications include industry $\times$ year, occupation $\times$ year, within occupation–industry income bin $\times$ year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Figure A.6: Technology exposure and worker job loss, by age or worker relative income (alternative measure based on low/high skill tasks)



Note: This figure shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text, except now we use ChatGPT to classify individual tasks as requiring either low or high specific preparation required, and we replace earnings growth with a proxy for involuntary job loss as the dependent variable; we allow these coefficients to vary by age group (panel A) or within occupation-industry income rank (panel B). The dependent variable is an indicator for whether or not a worker leaves their current employer in the next five years and also experiences an income growth rate beneath the 20th percentile for the year. We plot 95% confidence intervals from standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. All specifications include industry $\times$ year, occupation $\times$ year, within occupation–industry income bin $\times$ year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Figure A.7: Distribution of AI exposure across occupations



Note: The figure shows the cross-sectional distribution of the average AI exposure in terms of its labor-saving and labor-augmenting applications by broad occupation category. The square corresponds to the average exposure within each broad group; the vertical marks correspond to the 25th and 75th percentile of the average exposure of occupations within each broad occupation category; the length of the lines corresponds to the p90-p10 range. Broad occupation categories correspond to 2-digit SOC codes; occupations correspond to 6-digit SOC codes.

Table A.1: Validating our measure using ChatGPT4

ChatGPT Query	% Agree			
	5 **most** similar OCCs		5 **least** similar OCCs	
	R/N	L/H Exp	R/N	L/H Exp
	(1)	(2)	(3)	(4)
Distance to Routine Tasks / Low Exp: *"Here is the abstract of a patent: [Patent abstract here]. Here are some tasks: [DOT title here]. Do you think the technology mentioned above can perform some tasks mentioned above formerly performed by workers? Output yes or no, and your reasoning in one sentence."*	86%	84%	4%	21%
Distance to Non-Routine Tasks / High Exp: *"Here is the abstract of a patent: [Patent abstract here]. Here are some tasks: [DOT title here]. Do you think the patent mentioned above can increase the productivity of workers when performing some of the tasks mentioned above? Output yes or no, and your reasoning in one sentence."*	83%	83%	10%	7%

Note: In this table we randomly select 10 thousand breakthrough patents and compare ChatGPT assessments of occupational relatedness to patents with our patent-occupation textual similarity scores. For each patent we take the 5 most- and least-similar occupations based on $\rho^j(b, o)$ from (14) for routine/non-routine or low/high experience required tasks. We then provide ChatGPT a list of DOT occupational microtitles associated with the given 6-digit SOC occupation; for the five occupations with the most- and least-similar routine or low experience tasks, we then ask whether the given patent could perform some of these tasks instead of workers (query in top row); for the five occupations with the most- and least-similar non-routine or high experience tasks, we ask whether the given patent could improve worker productivity in performing these tasks (query in bottom row). In the first and third columns we report the average probability that ChatGPT responds affirmatively to the query for the most- and least-textually similar occupations, respectively, based off routine/non-routine tasks; we report the analogous probabilities instead using low/high experience required tasks in the second and fourth columns.

Table A.2: Summary Statistics: Census-CPS merged sample (worker-level data)

Variable	Mean	SD	5%	10%	25%	Median	75%	90%	95%	Observations
W2 Earnings	66,150	145,500	15,500	20,710	32,390	50,190	76,070	114,000	152,200	2,782,000
Age	40.9	7.4	29	31	35	41	47	51	53	2,782,000
Age, workers in bottom-25 income bin	40.0	7.6	29	30	33	40	46	51	53	632,000
Age, workers in top-5 income bin	43.2	6.8	31	33	38	44	49	52	53	109,000
Occupation-industry technology exposure (ξ)	0.647	0.971	0	0	0	0.232	0.866	2.028	2.922	1,495,000
Lifecycle-adjusted earnings growth, 3-years	-0.072	0.478	-0.973	-0.507	-0.129	0.008	0.129	0.312	0.472	2,773,000
Lifecycle-adjusted earnings growth, 5-years	-0.095	0.526	-1.116	-0.622	-0.174	0.002	0.142	0.337	0.507	2,596,000
Lifecycle-adjusted earnings growth, 10-years	-0.145	0.609	-1.363	-0.825	-0.280	-0.023	0.160	0.388	0.576	1,697,000
Male	0.542	0.498	0	0	0	1	1	1	1	2,782,000
Has four-year college degree	0.344	0.475	0	0	0	0	1	1	1	2,782,000
Involuntary exit proxy	0.1775	0.3821	0	0	0	0	0	1	1	2,147,000
5-year earnings growth $\mid$ Inv. exit = 1	-0.9447	0.6705	-2.55	-2.06	-1.229	-0.7033	-0.4407	-0.331	-0.2924	381,000
Exit firm within 1 year	0.175	0.380	0	0	0	0	0	1	1	2,676,000
Exit firm within 5 years	0.516	0.500	0	0	0	0	1	1	1	2,210,000
Union member	0.165	0.271	0	0	0	0	0	1	1	553,000
Industry unionization rate	0.163	0.163	0	0	0.031	0.111	0.250	0.436	0.475	2,398,000
Log revenues per worker	5.048	1.106	3.227	3.691	4.468	5.106	5.741	6.375	6.745	966,000

Note: The table reports summary statistics for our wage earnings data from the Census Detailed Earnings Record (DER)-CPS merged sample, which covers the 1981 to 2016 period. The sample includes all workers whose unique identifiers (PIK codes) can be matched between the DER and CPS data for CPS years between 1981 and 2016 and who satisfy labor force attachment sampling criteria. W2-Earnings are reported in terms of 2015 dollars. The occupation-industry technology exposure ξ is defined as in (15) from the main text. Patents are matched to industry of origination using information from the confidential Census SSL and LBD datasets. The variable "Has four-year college degree" denotes whether a given individual has completed a 4-year degree at the time they were observed in the CPS. The involuntary exit proxy equals 1 if a worker leaves the firm within the next 5 years and also has a 5-year earnings growth rate that is below the yearly 20th percentile. Workers are required to be between the ages of 25 and 55 to be included in the sample. Lifecycle-adjusted earnings growth rates follow Guvenen et al. (2014) and are constructed following (17) in the main text. Log revenues per worker come from the Longitudinal Business Database revenues file and have coverage from 1997 to the end of our sample. We winsorize earnings growth rates, as well as the summation component of worker technology exposure (defined in equation (15)), at the 1% level each year. Observation counts are rounded in accordance with Census disclosure rules. For more details on the construction of the CPS-DER matched sample and the linking of patents to industries, see appendix section B.7.

Table A.3: Technology Exposure and Worker Earnings Growth
(alternative measure based on low/high skill tasks)

	(1)	(2)	(3)	(4)
A. Horizon: 3 years of worker earnings growth				
Labor-saving Exposure ξ^L	-1.734	-1.638	-1.924	-1.975
	(-7.48)	(-6.93)	(-6.44)	(-6.33)
Labor-augmenting Exposure ξ^H	-0.982	-0.933	-1.229	-1.431
	(-3.61)	(-3.44)	(-3.25)	(-3.55)
B. Horizon: 5 years of worker earnings growth				
Labor-saving Exposure ξ^L	-1.938	-1.740	-2.269	-2.275
	(-7.47)	(-6.61)	(-6.85)	(-6.56)
Labor-augmenting Exposure ξ^H	-0.982	-0.880	-1.497	-1.718
	(-3.08)	(-2.83)	(-3.40)	(-3.7)
C. Horizon: 10 years of worker earnings growth				
Labor-saving Exposure ξ^L	-2.293	-1.902	-2.708	-2.617
	(-7.32)	(-5.99)	(-6.71)	(-6.23)
Labor-augmenting Exposure ξ^H	-1.039	-0.991	-1.840	-2.127
	(-2.92)	(-2.8)	(-3.54)	(-3.95)
Fixed Effects				
Industry (NAICS 4-digit)	Y	Y		
Occupation (SOCx)	Y		Y	
Ind × Year			Y	Y
Occ × Year		Y		Y
Prior Income Rank × Year	Y	Y	Y	Y

Note: Table shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 3, 5, or 10 years. We report t statistics corresponding to standard errors clustered at the occupation-industry (NAICS 4-digit) level in parentheses beneath the coefficient estimates. We scale the coefficients so that they correspond to a shift from the median to the 90th percentile of our technology exposure measures. Columns (1) to (4) report coefficient estimates under different combinations of fixed effects; all specifications include prior income rank × calendar year fixed effects. Prior income rank bins are based on workers' yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there are fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification when assigning income ranks. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th and 95th percentile; and, 95th percentile and above. All specifications also include controls listed under Table 1. See main text and notes to Table 1 for further details.

Table A.4: Worker technology exposure and skill
(alternative measure based on low/high skill tasks)

	Alternative Technology Exposure	
	Labor-Saving (ξ^L)	Labor-Augmenting (ξ^H)
A. Worker Age		
25-35 yo	-2.28	-0.63
	(-5.54)	(-1.23)
35-45 yo	-1.62	-2.02
	(-4.34)	(-4.21)
45-55 yo	-3.26	-1.95
	(-7.66)	(-3.98)
Older-Younger	-0.98	-1.32
	(-2.42)	(-3.39)
B. Income (relative to Ind × Occ peers)		
0–25th percentile	-2.87	-0.85
	(-6.91)	(-1.59)
25–50th percentile	-2.05	-0.93
	(-5.2)	(-1.88)
50–75th percentile	-2.03	-1.60
	(-5.43)	(-3.42)
75–95th percentile	-1.96	-2.90
	(-4.63)	(-5.74)
95–100th percentile	-2.45	-5.59
	(-3.7)	(-6.84)
Top-Bottom	0.42	-4.74
	(0.63)	(-5.61)
C. Education		
No College Education	-2.41	-1.27
	(-6.97)	(-2.82)
College Educated	-1.89	-2.10
	(-3.92)	(-4.42)
College-NoCollege	0.52	-0.83
	(1.31)	(-3.26)

Note: This table shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text, except now we use ChatGPT to classify individual tasks as requiring either low or high specific preparation required; we allow these coefficients to vary by age group (panel A); within occupation-industry income rank (panel B); or 4-year college graduate status (panel C). The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report t-statistics based on standard errors clustered at the occupation-industry level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. All specifications include industry × year, occupation × year, within occupation–industry income bin × year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Table A.5: Technology exposure and worker earnings growth by income rank, alternative income rankings

Income Rank	Baseline		Drop Recent Hires		2-yr Avg Income		Firm Adjusted		Residual Earnings: Age/Sex/CZ		Residual Earnings: Age/Sex/CZ/F		Residual Earnings: Age/CZ/Union	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)	
	ξ^R	ξ^N	ξ^R	ξ^N	ξ^R	ξ^N	ξ^R	ξ^N	ξ^R	ξ^N	ξ^R	ξ^N	ξ^R	ξ^N
0–25th percentile	-2.72	-0.82	-3.02	-0.68	-3.19	-0.90	-3.02	-0.69	-3.46	-0.52	-3.55	-0.06	-3.24	-0.99
	(0.38)	(0.55)	(0.41)	(0.57)	(0.36)	(0.51)	(0.4)	(0.55)	(0.4)	(0.57)	(0.42)	(0.55)	(0.74)	(0.97)
25–50th percentile	-1.96	-0.80	-1.59	-0.50	-1.68	-0.99	-1.91	-0.69	-2.04	-0.18	-1.90	-0.37	-2.47	-1.03
	(0.36)	(0.52)	(0.37)	(0.54)	(0.33)	(0.48)	(0.36)	(0.52)	(0.36)	(0.53)	(0.36)	(0.51)	(0.63)	(0.90)
50–75th percentile	-2.30	-1.11	-1.95	-0.81	-1.81	-1.43	-2.10	-0.93	-2.12	-1.04	-1.81	-0.94	-2.41	-1.00
	(0.34)	(0.49)	(0.34)	(0.52)	(0.32)	(0.47)	(0.35)	(0.51)	(0.34)	(0.51)	(0.34)	(0.52)	(0.61)	(0.89)
75–95th percentile	-2.58	-2.09	-2.28	-1.53	-2.27	-1.99	-2.09	-1.78	-2.47	-1.42	-2.28	-1.27	-2.90	-1.39
	(0.36)	(0.53)	(0.37)	(0.55)	(0.34)	(0.51)	(0.37)	(0.54)	(0.36)	(0.53)	(0.37)	(0.53)	(0.67)	(0.95)
95–100th percentile	-3.12	-4.74	-2.77	-4.05	-2.38	-4.43	-3.27	-3.68	-3.33	-4.62	-3.31	-4.21	-2.27	-5.41
	(0.64)	(0.84)	(0.66)	(0.84)	(0.57)	(0.85)	(0.66)	(0.80)	(1.69)	(1.44)	(0.47)	(0.67)	(0.80)	(0.86)
Top-Bottom (p-val)	0.54	0.00	0.71	0.00	0.14	0.00	0.82	0.00	0.53	0.00	0.74	0.00	0.53	0.00

Note: Table shows the estimated slope coefficients γ and δ (times 100) from equation (16) in the main text, where these coefficients vary with worker earnings rank. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the occupation-industry (NAICS 4-digit) level beneath coefficient estimates, and normalize the coefficients to correspond to a shift from the median to the 90th percentile. Columns (1) through (7) use different methods for ranking workers based on earnings. Column (1) represents our baseline method of sorting workers each year within occupation–industry. Column (2) uses our baseline income sort but drops workers who were hired within the last year. Column (3) ranks workers on average earnings over the past two years. Column (4) adjusts workers' earnings for average firm wages by subtracting off the log average wage of a worker's employer in the Longitudinal Business Database from the worker's log wage. Column (5) residualizes log earnings with respect to yearly fixed effects for occupation–industry; commuting zone; and 10-year age bin (25-35, 35-45, 45-55) interacted with gender. Column (6) residualizes income with respect to the same characteristics as column (5), plus year-specific Longitudinal Business Database firm wage decile bins; Column (7) restricts to subsample of workers who are in-universe for the CPS union membership question, and residualizes log earnings with respect to yearly fixed effects for occupation, industry, worker union status, and commuting zone. The bottom panel of the table reports the p-value of corresponding to a two-sided test of coefficient equality between the top and bottom income bin coefficients in each column. All specifications include industry × year, occupation × year, within occupation–industry income bin × year fixed effects, dummies at the level of coefficient interaction, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Table A.6: Technology exposure and worker outcomes, with spillovers

	Baseline			Alternative Measure		
	ξ^R	ξ^N	$\bar{\xi}$	ξ^L	ξ^H	$\bar{\xi}$
A. All Workers						
homogeneous Coeff	-0.030	-0.017	0.025	-0.031	-0.020	0.027
	(-7.84)	(-2.75)	(5.41)	(-7.39)	(-3.36)	(6.14)
B. Coefficients vary by Relative Income						
homogeneous Coeff			0.025			0.027
			(5.42)			(6.12)
0–25th percentile	-0.034	-0.011		-0.038	-0.010	
	(-7.57)	(-1.67)		(-7.78)	(-1.62)	
25–50th percentile	-0.026	-0.013		-0.029	-0.013	
	(-6.25)	(-1.91)		(-6.45)	(-2.07)	
50–75th percentile	-0.028	-0.016		-0.028	-0.019	
	(-7.11)	(-2.5)		(-6.47)	(-3.26)	
75–95th percentile	-0.030	-0.025		-0.027	-0.031	
	(-7.17)	(-3.77)		(-5.66)	(-5.07)	
95–100th percentile	-0.036	-0.050		-0.032	-0.057	
	(-5.55)	(-5.47)		(-4.66)	(-6.5)	
Top-Bottom	-0.002	-0.039		0.006	-0.046	
	(-0.29)	(-4.50)		(0.95)	(-5.58)	

Note: This table shows the estimated slope coefficients α, γ and δ (times 100) from equation (25) in the main text; in panel A we estimate homogeneous coefficients for γ and δ on individual-level exposure measures ξ^R and ξ^N, respectively, while we allow these coefficients to vary by an individuals' within-occupation-industry earnings rank in panel B. The specification in the first 3 columns corresponds to our measures that use the baseline routine/non-routine classification of tasks, while the last three columns use our alternative designation of tasks into low or high amounts of skill required. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report t-statistics from standard errors clustered at the occupation-industry level beneath coefficient estimates, and we normalize the coefficients to correspond to a shift from the median to the 90th percentile for individual-level exposures ξ^R and ξ^N (similarly for ξ^L and ξ^H) and a standard deviation increase for $\bar{\xi}$. Since we use 4-digit NAICS level average innovation exposure $\bar{\xi}$, we replace the usual 4-digit NAICS $\times$ year fixed effects with 2-digith NAICS $\times$ year fixed effects. Specifications also control for the log of industry employment and include occupation $\times$ year, within occupation–industry income bin $\times$ year fixed effects, and controls listed under Table 1. See main text and notes to Table 1 for further details.

Table A.7: Summary of Targeted Moments

#	Parameter Combination	Description
1	$\frac{\psi-\nu_R}{\nu_R+\zeta_R}\frac{(\nu_R+\zeta_R)\,\kappa_R}{\psi+\zeta_R+(\nu_R+\zeta_R)\,\kappa_R}\tilde{\sigma}_R$	worker earnings to ξ^R, homogeneous
2	$\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}-\beta\right]\tilde{\sigma}_N$	worker earnings to ξ^N, homogeneous
3	$\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}-\beta+\sigma\omega\frac{f(Q_{0.25})-f(Q_0)}{\Phi(Q_{0.25})-\Phi(Q_0)}\right]\tilde{\sigma}_N$	worker earnings to ξ^N, 0%-25%
4	$\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}-\beta+\sigma\omega\frac{f(Q_{0.5})-f(Q_{0.25})}{\Phi(Q_{0.5})-\Phi(Q_{0.25})}\right]\tilde{\sigma}_N$	worker earnings to ξ^N, 25%-50%
5	$\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}-\beta+\sigma\omega\frac{f(Q_{0.75})-f(Q_{0.5})}{\Phi(Q_{0.75})-\Phi(Q_{0.5})}\right]\tilde{\sigma}_N$	worker earnings to ξ^N, 50%-75%
6	$\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}-\beta+\sigma\omega\frac{f(Q_{0.95})-f(Q_{0.75})}{\Phi(Q_{0.95})-\Phi(Q_{0.75})}\right]\tilde{\sigma}_N$	worker earnings to ξ^N, 75%-95%
7	$\left[\frac{\psi-\nu_N}{\nu_N+\zeta_N}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}-\beta+\sigma\omega\frac{f(Q_1)-f(Q_{0.95})}{\Phi(Q_1)-\Phi(Q_{0.95})}\right]\tilde{\sigma}_N$	worker earnings to ξ^N, 95%-100%
8	$\tilde{\beta}_1$ Given by (A.71)	worker earnings to ξ^R, regression with $\bar{\xi}$, homogeneous
9	$\tilde{\beta}_2$ Given by (A.71)	worker earnings to ξ^N, regression with $\bar{\xi}$, homogeneous
10	$\tilde{\delta}$ Given by (A.71)	worker earnings to $\bar{\xi}$, homogeneous
11	$\frac{1}{1+\gamma^2}\frac{1}{1+\chi\,\epsilon_c}\frac{(\nu_R+\zeta_R)\,\kappa_R}{\psi+\zeta_R+(\nu_R+\zeta_R)\,\kappa_R}\tilde{\sigma}_{\text{comp}}$ (imposing $\Gamma_R=\Gamma_N$ here)	industry prod to $\bar{\xi}$
12	$\frac{1}{1+\gamma^2}\frac{(\nu_R+\zeta_R)\,\kappa_R}{\psi+\zeta_R+(\nu_R+\zeta_R)\,\kappa_R}\left[(1-\nu_R)\frac{\psi+\zeta_R}{\nu_R+\zeta_R}+(\psi-1)(1-\frac{LS}{LS_R})+\frac{LS}{LS_R}\frac{\vartheta}{1+\chi\epsilon_c}\right]\tilde{\sigma}_{\bar{R}}$	industry labor share to $\bar{\xi^R}$
13	$\frac{1}{1+\gamma^2}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}\left[(1-\nu_N)\frac{\psi+\zeta_N}{\nu_N+\zeta_N}+(\psi-1)(1-\frac{LS}{LS_N})+\frac{LS}{LS_N}\frac{\vartheta}{1+\chi\epsilon_c}\right]\tilde{\sigma}_{\bar{N}}$	industry labor share to $\bar{\xi^N}$
14	$(\psi-\nu_R)\frac{\zeta_R}{\nu_R+\zeta_R}\frac{(\nu_R+\zeta_R)\,\kappa_R}{\psi+\zeta_R+(\nu_R+\zeta_R)\,\kappa_R}\tilde{\sigma}_{R_o}$	occupation employment to ξ^R
15	$(\psi-\nu_N)\frac{\zeta_N}{\nu_N+\zeta_N}\frac{(\nu_N+\zeta_N)\,\kappa_N}{\psi+\zeta_N+(\nu_N+\zeta_N)\,\kappa_N}\tilde{\sigma}_{N_o}$	occupation employment to ξ^N
16	$[\theta(\kappa_R+1)+(1-\theta)(\kappa_N+1)]^{-1}$	labor share

Note: This table displays close-form model-implied expressions for all moments we target in our GMM estimation. See appendix section A for model details, including subsection A.8 for an in-depth discussion of model estimation.

Table A.8: Most Exposed Occupations to AI as automation

Occupation	6-digit SOC	Automation	Complementary	Skill Displacement	Total
Tellers	433071	-12.6	2.6	-3.2	-13.2
Word Processors and Typists	439022	-11.7	3.7	-4.6	-12.6
Data Entry Keyers	439021	-11.6	3.7	-4.7	-12.6
Bookkeeping, Accounting, and Auditing Clerks	433031	-11.6	3.1	-3.9	-12.4
Food Cooking Machine Operators and Tenders	513093	-11.1	3.3	-4.1	-11.9
Mail Clerks and Mail Machine Operators, Except Postal Service	439051	-11.0	2.8	-3.5	-11.7
Switchboard Operators, Including Answering Service	432011	-10.7	4.9	-6.1	-12.0
Postal Service Clerks	435051	-10.4	5.2	-6.6	-11.8
Graders and Sorters, Agricultural Products	452041	-10.1	3.2	-4.1	-11.0
Adhesive Bonding Machine Operators and Tenders	519191	-10.1	1.2	-1.4	-10.4
Other Information And Records Clerks	434YYY	-9.9	3.8	-4.7	-10.9
Postal Service Mail Carriers	435052	-9.9	0.6	-0.8	-10.0
Photographic Process Workers and Processing Machine Operators	519151	-9.8	3.2	-4.0	-10.6
Bus Drivers, Transit And Intercity	533052	-9.8	1.3	-1.7	-10.1
Billing and Posting Clerks	433021	-9.8	5.6	-7.0	-11.2
Hotel, Motel, and Resort Desk Clerks	434081	-9.6	3.7	-4.7	-10.6
Payroll and Timekeeping Clerks	433051	-9.6	5.7	-7.1	-11.1
Helpers–Production Workers	519198	-9.5	4.3	-5.5	-10.6
Packaging and Filling Machine Operators and Tenders	519111	-9.5	4.6	-5.8	-10.7
Postal Service Mail Sorters, Processors, and Processing Machine Operators	435053	-9.5	2.1	-2.6	-10.0
Paper Goods Machine Setters, Operators, and Tenders	519196	-9.4	1.4	-1.8	-9.8
Office Clerks, General	439061	-9.4	5.0	-6.3	-10.6
Shipping, receiving, and inventory clerks	435071	-9.3	5.0	-6.2	-10.6
Miscellaneous Production Workers, Including Equipment Operators and Tenders	5191XX	-9.3	2.5	-3.2	-10.0
Prepress Technicians and Workers	515111	-9.3	3.2	-4.0	-10.1

Note: In this table we show the occupations with highest exposure to labor-complementing potential of artificial intelligence technologies using our GMM model parameter estimates, where we use GPT4 to classify descriptions occupation tasks from August 2023 edition of ONET into being complemented or substituted by AI and exposed or not exposed to AI-related cost improvements. The labor-substituting exposure is defined in appendix equation (A.83). We also report the different components of AI exposure-related earnings changes for these occupations, as in Table 6. See main text and appendix A (especially subsection A.9) for further details.

Table A.9: Most Exposed Occupations to AI as complement

Occupation	6-digit SOC	Automation	Complementary	Skill Displacement	Total
Insurance Underwriters	132053	0.0	13.3	-16.7	-3.4
Medical Transcriptionists	319094	-1.9	12.5	-15.6	-5.1
Customer Service Representatives	434051	-3.1	12.4	-15.6	-6.3
Personal Financial Advisors	132052	0.0	12.3	-15.5	-3.1
Budget Analysts	132031	0.0	11.7	-14.7	-3.0
Loan Interviewers and Clerks	434131	-4.8	11.5	-14.4	-7.8
Retail Salespersons	412031	-3.2	11.4	-14.3	-6.1
Market Research Analysts and Marketing Specialists	131161	-0.7	10.9	-13.7	-3.5
Human Resources Assistants, Except Payroll and Timekeeping	434161	-3.8	10.9	-13.7	-6.6
Property appraisers and assessors	132020	-1.6	10.9	-13.7	-4.4
Telemarketers	419041	-6.1	10.8	-13.6	-8.8
Cost Estimators	131051	-1.0	10.8	-13.5	-3.8
Court, Municipal, and License Clerks	434031	-4.4	10.6	-13.3	-7.1
Court Reporters And Simultaneous Captioners	273092	-5.3	10.5	-13.2	-8.0
Advertising Sales Agents	413011	0.0	10.3	-13.0	-2.6
Proofreaders and Copy Markers	439081	-5.3	10.2	-12.8	-7.9
Financial and Investment Analysts	132051	-0.6	10.0	-12.6	-3.1
Surveying and Mapping Technicians	173031	-3.4	10.0	-12.5	-5.9
Credit Counselors and Loan Officers	132070	-2.8	9.9	-12.4	-5.3
Wholesale and Retail Buyers, Except Farm Products	131022	-2.9	9.9	-12.4	-5.5
Other Office And Administrative Support Workers	439XXX	-5.6	9.5	-11.9	-8.1
Interviewers, Except Eligibility and Loan	434111	-2.8	9.5	-11.9	-5.3
Door-to-Door Sales Workers, News and Street Vendors, and Related Workers	419091	-2.4	9.4	-11.8	-4.8
Project Management Specialists	131082	0.0	9.3	-11.7	-2.4
First-Line Supervisors of Non-Retail Sales	411012	-1.6	9.3	-11.6	-4.0

Note: In this table we show the occupations with highest exposure to labor-substituting potential of artificial intelligence technologies using our GMM model parameter estimates, where we use GPT4 to classify descriptions of occupation tasks from August 2023 edition of ONET into being complemented or substituted by AI and exposed or not exposed to AI-related cost improvements. The labor-complementing exposure is defined in appendix equation (A.84). We also report the different components of AI exposure-related earnings changes for these occupations, as in Table 6. See main text and appendix A (especially subsection A.9) for further details.