HUBEN LIU
Northwestern University
Kellogg School of Management

LAWRENCE D. W. SCHMIDT
MIT Sloan School of Management

DIMITRIS PAPANIKOLAOU
Northwestern University
Kellogg School of Management

BRYAN SEEGMILLER
Northwestern University
Kellogg School of Management

Technology and Labor Markets: Past, Present, and Future; Evidence from Two Centuries of Innovation

ABSTRACT We use recent advances in natural language processing and large language models to construct novel measures of technology exposure for workers that span almost two centuries. Combining our measures with census data on occupation employment, we show that technological progress over the twentieth century has led to economically meaningful shifts in labor demand across occupations: It has consistently increased demand for occupations with higher education requirements, occupations that pay higher wages, and occupations with a greater fraction of female workers. Using these insights and a calibrated model, we then explore different scenarios for how advances in artificial intelligence (AI) are likely to affect employment trends in the medium run. The model predicts a reversal of past trends, with AI favoring occupations that are lower-educated, lower-paid, and more male-dominated.

Conflict of Interest Disclosure: This paper received financial support from the Financial Institutions and Markets Research Center and the Asset Management Practicum. The Brookings Institution is committed to quality, independence, and impact. Brookings is supported by a diverse array of funders. In line with the Brookings values and policies, each Brookings publication represents the sole views of its author(s).

Brookings Papers on Economic Activity, Fall 2025: 215–272

Economists and workers alike have long worried about the prospect of technological displacement of labor.[1] Recent advances in artificial intelligence (AI) have reignited the perennial concern that technology will automate away most tasks performed by workers, leading to large declines in labor demand, depressed wages, and diminished job opportunities for workers. Yet systematic evidence on how technological advances shape labor demand over the long run remains limited. The difficulty is that even when technologies are labor-saving, their net impact is ambiguous: They directly substitute for labor in some tasks, but may also raise overall productivity, induce reallocation toward complementary tasks, and increase aggregate labor demand.

Our goal is twofold. First, we aim to tease apart these forces and understand the role that technological progress has played in shaping the demand for labor over the long run. Overall, we find that technological innovation has led to economically meaningful shifts in labor demand across occupations: It has consistently increased demand for occupations with higher education requirements, occupations that pay higher wages, and occupations with a greater fraction of female workers. Second, we explore the extent to which the experience over the last two centuries is informative about the role that advances in AI are likely to play in shaping the composition of the labor force in the medium run. In sharp contrast to the past two centuries, our framework suggests that AI—by substituting primarily for cognitive tasks—will shift relative demand toward occupations with lower education, lower wages, and a greater share of male workers.

Our starting point is a simple theoretical model based on Hampole and others (2025) that links task-specific technological advances to overall labor demand for an occupation. The model nests both direct and indirect channels. A technology that improves a task-specific form of capital substitutes for labor in that task. If it applies broadly across all tasks of an occupation—

1. In ca. 350 BCE, Aristotle wrote: "[If] the shuttle would weave and the plectrum touch the lyre without a hand to guide them, chief workmen would not want servants, nor masters slaves" (Aristotle 1885, p. 6). In 1811, skilled weavers and textile workers (the Luddites) worried that mechanizing manufacturing (and the unskilled laborers operating the new looms) would rob them of their means of income. In 1930, Keynes worried about technological unemployment: "We are being afflicted with a new disease of . . . technological unemployment . . . due to our discovery of means of economising the use of labour outrunning the pace at which we can find new uses for labour" (Keynes 1978, p. 325). More recently, a McKinsey report estimated that between 400 million and 800 million jobs could be lost worldwide due to new technologies by the year 2030 (Manyika and others 2017).

as with the automatic telephone switching system that displaced operators—the result is a sharp reduction in labor demand. If it applies only to a narrow set of tasks—say, a tool that automates expense reporting for academic economists—the negative impact on labor demand is mitigated, or possibly reversed. Specifically, because workers optimally reallocate time across tasks, automation in one task has ripple effects: It frees up labor for other tasks within the occupation, and it can raise overall productivity in ways that boost labor demand in other, not directly affected, occupations.

The model implies that the labor market impact of a specific technology can be summarized by three statistics. First, the *mean* exposure of an occupation's tasks to technology improvements is, in general, negatively related to demand for that occupation. Thus, a moderate improvement in a technology, which is related to all the tasks of a particular occupation will lower demand—even a basic tractor will reduce demand for farm labor. Second, the degree to which the technology exposure is *concentrated* in a subset of tasks allows workers to reallocate their effort and thus increases their productivity and the labor demand for their occupation. Third, the degree of productivity improvements as a result of these technological advances directly determines the shift in labor demand across all indirectly affected occupations.

Measuring workers' exposure to technology presents two core challenges: identifying technological improvements and determining the extent to which they substitute for specific tasks. We begin by using patent documents as our measure of innovation. While not all inventions are patented, the patent record provides a consistent proxy for technological change dating back to the mid-nineteenth century. Mapping technology to tasks is more subtle. Occupation-specific tasks evolve over time as the structure of work changes. To address this, we use a state-of-the-art large language model (LLM) with web search capabilities to generate comprehensive task descriptions for US Census occupations in each decade from 1850 to 2010. We then quantify exposure using modern natural language processing techniques: Both patent text and task descriptions are embedded as numerical vectors, and their semantic similarity is computed via cosine distance. A task is classified as exposed to a new technology if its vector representation closely aligns with that of contemporaneous patents. This approach allows us to track the evolving overlap between new technologies and the tasks performed by workers over time.

Naturally, the direction of technological innovation can be endogenous to the current state of the market for specific types of workers. To strengthen our interpretation of the findings as identifying the causal effect of technology on labor demand, we develop a shift-share identification strategy inspired by Acemoglu, Akcigit, and Kerr (2016). Specifically, our identification strategy leverages the degree to which breakthrough technological advances, according to Kelly and others (2021), in upstream technologies diffuses to innovations in downstream, labor-saving technologies. For example, our instrument leverages the proliferation of breakthrough technologies occurring in the upstream cooperative patent classification (CPC) technology class G11 (Information Storage), which led to downstream improvements in the technology class G06 (Computing, Calculating, and Counting) during the 1980s and 1990s; around the turn of the twentieth century, our instrumental variables (IV) approach predicts that upstream breakthroughs in technology class B61 (Railways) resulted in downstream innovations in technology class B60 (Vehicles in general).

Consistent with the model, we find that the average exposure of an occupation's tasks to the technologies developed in a given period is significantly negatively related to subsequent employment growth. However, for a given level of mean exposure, labor demand for that occupation increases if the technology exposure is concentrated in a subset of tasks. Examining how these results vary across different periods, we find that our results are broadly consistent across periods, with the exception of the 1880–1920 period. This early period coincides with a number of changes in the definitions of occupations, a shift that the census occupation classifications fail to adequately capture. Thus, in the remainder of the paper we focus on the post-1900 period. Doing so allows us to also control for differential industry trends (the census starts collecting information on industry in 1910). Our results remain similar if we include the interaction of industry and decade year effects—in which case we are comparing employment growth in differentially exposed occupations within a particular industry. Last, using the census industry assignments starting in 1910, our estimates reveal a positive relation between increases in the rate of patents that are relevant for a particular industry and employment across all occupations in that industry.

Taken together, our findings paint a fuller picture of how technological change has shaped the relative labor demand for different occupations over the last century or more. On the one hand, technology has directly substituted for specific worker tasks, which ceteris paribus has decreased labor demand for these tasks. However, this does not necessarily imply that the demand for labor decreased, due to the presence of these two quantitatively relevant offsetting forces: Labor-saving technologies allowed workers to direct their effort to tasks not substituted by technology, while the resulting productivity improvements increased the overall demand for labor for affected occupations relative to others.

That said, the relative importance of these forces varies across occupations. In particular, consistent with the prevailing view on job polarization (Autor and Dorn 2013; Autor, Katz, and Kearney 2006; Goos, Manning, and Salomons 2014), we find that occupations at the middle of the skill (income) distribution have been significantly more exposed to technology in the post-1980 period than workers at either the top or the bottom of the distribution. Importantly, however, our results indicate that this pattern of technology-induced job polarization likely started earlier—possibly to the middle and even early twentieth century. This is consistent with Bárány and Siegel (2018), who argue that labor market polarization can be dated to at least the post-1950 period, not just post-1980. Moreover, we even find some evidence of technology-related polarization slightly earlier, going back to the 1910–1960 subsample of our data. In sum, we find that job polarization is not just a post-1980 story. It has been going on for decades, and technology accounts for a substantial share. But consistent with prior work, the phenomenon intensifies after 1980, and our technology measures play a larger role in that period.

In addition, we sort occupations into employment-weighted quintiles based on their average educational attainment or share of workers who are female. Across the whole sample period and within each subperiod we analyze, we find a consistent pattern of technology-induced decline in the employment shares of less educated occupations and also of male-intensive occupations. Thus our measures of occupational technological exposure are consistent with the pattern of rising returns to education in the labor market (Katz and Murphy 1992; Card and Lemieux 2001; Goldin and Katz 2008) and its relation to skill-biased technological change (Berman, Bound, and Machin 1998; Krusell and others 2000). Our results on technology-induced employment expansion in female-heavy occupations in turn may also speak to the long-run decline in male labor force participation, which dates back at least to the mid-twentieth century (Parsons 1980).

A key advantage of our measure is that it also includes time series variation in the direction of innovation over time. Examining the time series, we see that, prior to 1980, innovation was consistently associated with manual physical tasks; by contrast, the innovations of the late twentieth and early twenty-first century have become relatively more related to cognitive tasks. This pattern is partly driven by the increased prevalence of breakthrough patents related to computers and electronics. Last, occupations that are associated with interpersonal tasks have consistently low exposures to innovation throughout the entire sample period.

A natural question is whether the substitutability between new technologies and worker tasks is stable across task types and over time. To address this, we decompose our mean task exposure measure into manual, cognitive, and interpersonal components and relate each to subsequent occupation-level employment changes. Three patterns emerge. First, exposure of manual tasks is consistently associated with employment declines, indicating that technological progress has persistently substituted for manual effort. Second, exposure of cognitive tasks shows a time-varying relation: Before the information and communications technology (ICT) revolution, it is associated with employment gains, but after the 1960s the effect turns negative. This shift suggests that earlier innovations complemented cognitive work, while more recent advances—especially in computing and software—have substituted for it. Finally, exposure of interpersonal tasks shows no systematic link to employment changes, implying that social skills may provide some insurance against labor-saving technologies, consistent with their rising importance in the labor market documented by Deming (2017).

We also exploit variation across worker cohorts to study how the link between employment growth and technology exposure differs by age. The employment changes we document are not driven solely by younger cohorts avoiding entry into exposed occupations. Instead, we find a strong negative relationship between mean exposure and employment growth among incumbents, tracking each age group forward as it ages. In specifications without the exposure concentration measure, the magnitude of the exposure coefficients increases steadily from the youngest to the oldest cohorts, consistent with human capital having a vintage-specific component. When we include concentration, however, the pattern shifts: Concentration coefficients are especially large for the youngest incumbents, suggesting that the productivity gains from task reallocation accrue disproportionately to younger workers.

In sum, our empirical results largely validate the model's predictions regarding the impact of technology exposure on labor demand. In the final part of the paper we use the structure of the model, with some additional assumptions, to explore the impact of advances in AI on relative labor demand across occupations over the medium run. Overall, a robust pattern emerges: We expect that AI advances will partially reverse the trends in relative employment growth induced by technology during the twentieth century. That is, we expect that AI will increase the relative demand for occupations with lower education requirements, lower pay, and lower share of female workers. These cross-sectional predictions are obtained under the assumption that AI is likely to automate certain cognitive tasks performed by workers, which do not require significant prior experience—in much the same way that mechanization in the twentieth century substituted for manual tasks performed by workers.

Our analysis connects to a large literature on labor-substituting technological change. One influential strand emphasizes the role of automation in displacing routine tasks, showing that occupations with higher routine task intensity have been more exposed to recent labor-saving advances (Autor, Katz, and Kearney 2006; Acemoglu and Autor 2011; Goos, Manning, and Salomons 2014). A second strand develops direct measures of labor-saving technologies and quantifies their labor market consequences.[2] Our contribution is to bring a long-run perspective: Rather than focusing on recent decades or specific technologies, we trace how successive waves of innovation since the mid-nineteenth century have reshaped labor demand across occupations.

Related work by Autor and Thompson (2025) also emphasizes how labor-substituting technological change can have nuanced effects depending on which specific tasks within a job are affected. Their model highlights how automation-induced changes in expertise requirements within a job operate like an occupational supply shift: Jobs that remove relatively nonexpert tasks due to automation may see both falling employment and rising wages, because the remaining nonautomated tasks require workers to maintain a higher level of expertise. Our mechanism instead emphasizes how the distribution of task exposure to technological change may mediate the size of the labor demand shift: Comparing two jobs facing the same average amount of labor-saving technological exposure across all their tasks, we expect the job with technology exposure that is more concentrated in a subset of tasks to experience better demand outcomes. This is because the ability to reallocate effort to unaffected tasks mitigates the negative direct effects of average exposure to technological substitution.

Closest to our work is Hampole and others (2025), who leverage detailed data on worker resumes and firm job postings to examine the impact of

2. An incomplete list includes Webb (2020); Jiang, Tang, and others (2025); Acemoglu and Restrepo (2022); Humlum (2021); Dauth and others (2021); Koch, Manuylov, and Smolka (2021); de Souza and Li (2023); Kogan and others (2023); Mann and Püttmann (2023); Jiang, Park, and others (2025); Hémous and others (2025); and Aghion and others (2020).

AI on labor demand during the 2010–2020 period. Using their model as a guide, we leverage data on patents to construct measures of the mean and concentration of technology exposure at the task level over the last two centuries. Given the long time span of our analysis, the granularity of our exposure measures is significantly lower—we can only construct our exposure measures at the occupation level. Nevertheless, we reach some similar conclusions: Occupations with high mean exposure to labor-saving technologies experience declines in labor demand, while the degree to which technology exposure is concentrated in a subset of tasks has an offsetting effect. Unlike Hampole and others (2025), we cannot observe technology adoption at the firm level, which limits our ability to estimate the degree of productivity spillovers in the data.

I. Theoretical Framework

To guide measurement, we first begin with a simple model based on Hampole and others (2025). A central implication of their framework is that the effect of technological change on labor demand depends on how capital-augmenting shocks are distributed across the tasks that define an occupation. When improvements are broad-based—raising capital productivity across most tasks—capital substitutes directly for labor, leading to a decline in occupational labor demand. By contrast, when advances are uneven and concentrated in a subset of tasks, workers shift effort toward the remaining tasks, raising their marginal productivity and potentially increasing labor demand overall.

I.A. Setup

A single final consumption good (the numeraire) is produced using a nested constant elasticity of substitution (CES) production function. In the outer nest, aggregate output $\overline{Y}$ is a CES composite of the output Y_I of a continuum of different industries indexed by I,

$$\overline{Y} = \left(\int_I \alpha_I^{\frac{1}{\theta}} Y_I^{\frac{\theta-1}{\theta}} dI \right)^{\frac{\theta}{\theta-1}}, \tag{1}$$

where θ captures the elasticity of substitution across industries and α_I is a weight capturing an industry's importance in the final good. Here, α_I can represent the level of industry total factor productivity or the number of products produced by the industry; see Hampole and others (2025) for more details.

Each industry produces its output Y_I by combining the output of many occupations,

$$(2) \qquad Y_I = \left(\int_O \alpha(o, I)^{\frac{1}{\chi}} Y(o, I)^{\frac{\chi-1}{\chi}} \right)^{\frac{\chi}{\chi-1}}.$$

The parameter χ governs the elasticity of substitution across different occupations within an industry, and $\alpha(o, I)$ is a weight capturing the importance of occupation o in industry I output. Relative to Hampole and others (2025), we simplify by assuming that firms are perfectly competitive within industries. Hence, all firms price at marginal cost and earn zero profits in equilibrium.

Workers in occupation o employed in industry I produce output $Y(o, I)$ as a CES composite of different tasks,

$$(3) \qquad Y(o, I) = \left(\sum_{j \in \mathcal{J}(o, I)} \alpha(j)^{\frac{1}{\psi}} y(j)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}},$$

where the relevant subset of tasks $\mathcal{J}(o, I)$ is occupation-specific. The CES parameter ψ governs the elasticity of substitution across tasks within a given job, controlling the extent to which tasks are complements ($\psi < 1$) or substitutes ($\psi > 1$) in production of $Y(o, I)$. We economize on notation at the task level by suppressing the industry and occupation subscripts unless needed. As above, the weight $\alpha(j)$, which may vary by industry, occupation, and task, captures the importance of task j in production.

Each task j in job (o) is produced by a labor input $l(j)$ and a capital input $k(j)$,

$$(4) \qquad y(j) = \left(\gamma_j l(j)^{\frac{\nu-1}{\nu}} + (1 - \gamma_j) k(j)^{\frac{\nu-1}{\nu}} \right)^{\frac{\nu}{\nu-1}}.$$

Here, ν captures the elasticity of substitution between capital $k(j)$ and labor $l(j)$. In discussing comparative statics obtained from the model, we assume that $\nu > \psi$. This assumption implies that reductions in the user cost of $k(j)$ specific to task j will likely be labor-saving.

As in Acemoglu and Restrepo (2022), $q(j)$ captures the quality-adjusted price of capital $k(j)$ that is specific to task j—that is, the rate at which the final good can be transformed into capital specific to task j. Given our objective of capturing long-run shifts in the economy, capital is assumed to fully depreciate after production, and part of the final good is used as an intermediate input in production of capital. Following Hampole and others (2025) and Kogan and others (2023), innovations affect the real economy by reducing the quality-adjusted price of capital $q(j)$,

$$\Delta \log q(j) = -\varepsilon(j). \tag{5}$$

Each technological improvement is potentially applicable to several tasks within a job—defined as an occupation-industry pair. A given technological improvement that is applicable to job (o) can therefore be represented as a job-specific vector ε of weakly positive random variables. If $\varepsilon(j) > 0$, that implies the arrival of an improved (or cheaper) labor-saving technology that is specific to task j. For now, we are completely agnostic about the joint distribution of these technology improvements.

The last piece is labor supply. The first-order condition for labor supply at the occupation-industry level satisfies

$$N(o, I) = \alpha_I \alpha_I(o)\left(W(o, I)\right)^{\zeta}, \tag{6}$$

where $W(o, I)$ is the equilibrium wage offered to workers in industry I and occupation o.

In addition to choosing how many efficiency units of labor to allocate to each occupation-industry cell, the Hampole and others (2025) framework allows workers to optimally choose the fraction of time to allocate to each task in order to maximize total earnings. The effective supply of labor by workers in task j is equal to

$$l(j) = \alpha(j)^{\beta} h(j)^{1-\beta}, \tag{7}$$

where $\beta \in (0, 1)$ captures the degree of decreasing returns to effort at the task level. As $\beta \rightarrow 1$, the allocation of efficiency units across tasks is fixed at the exogenous level $\alpha(j)$. By contrast, as $\beta \rightarrow 0$, workers can frictionlessly reallocate time across tasks. Lowering β increases the flexibility of workers to respond to changes in their productivity across tasks by reallocating time. For each efficiency unit of labor, we normalize the total number of hours a worker can supply across all tasks, which is equal to one. Given the above, the optimal hours supply of a worker in job (o, I) to task j is equal to

(8) $$h(j) = \frac{\alpha(j) w(j)^{\frac{1}{\beta}}}{\sum_{k \in \mathcal{J}(o,I)} \alpha(k) w(k)^{\frac{1}{\beta}}},$$

where $w(j)$ is the occupation-specific wage in task j. Thus, incorporating the optimal hours choice in equation (8), a worker's total earnings in job (o) per efficiency unit equals

(9) $$W(o,I) \equiv \sum_{j \in \mathcal{J}(o,I)} \alpha(j)^{\beta} h(j)^{1-\beta} w(j) = \left(\sum_{j \in \mathcal{J}(o,I)} \alpha(j) w(j)^{\frac{1}{\beta}} \right)^{\beta},$$

which depends on the worker's allocation of time and the (job-specific) task prices $w(j)$.

I.B. Model Implications

With our model in hand, we can analyze the impact of a shift in ε on equilibrium earnings and employment. Approximating around the symmetric equilibrium in which the labor share is constant across tasks, Hampole and others (2025) derive a simple linear equation for employment growth due to changes in technology,

(10) $$\Delta \log N(o,I) \approx \underbrace{\zeta \eta_m m(\varepsilon) + \frac{\zeta}{2\beta} \eta_o^2 C(\varepsilon)}_{\text{Direct effects}} + \underbrace{\Delta \log \alpha_I + \zeta \eta_z \Delta_\varepsilon \log Z_I}_{\text{Industry spillovers}} + \underbrace{\frac{\zeta \eta_z}{\theta - \chi} \Delta_\varepsilon \log \overline{\Omega}}_{\text{Aggregate spillovers}}.$$

Equation (10) serves as the foundation of our empirical analysis. We next discuss its key components.

The first two terms in equation (10) capture the direct effect of technology improvements on labor demand, which depends on two key sufficient statistics. The first statistic is

(11) $$m(\varepsilon) \equiv \sum_{j \in \mathcal{J}(o,I)} \frac{\alpha(j)}{\sum_{k \in \mathcal{J}(o,I)} \alpha(k)} \varepsilon(j),$$

that is, the task importance–weighted mean improvement of the technology across all tasks within the industry-occupation cell. The impact of the occupation's mean exposure, equation (11), on labor demand depends on the elasticity of labor demand ζ and

$$\eta_m \equiv -\frac{s_k\left(\nu-\chi\right)}{\zeta+\nu s_k+\chi\left(1-s_k\right)}, \tag{12}$$

where s_k is the capital share, which is assumed to be equal across all tasks.

The sign of η_m depends on whether the elasticity of substitution between capital and labor ν exceeds the elasticity of substitution χ across occupations within industries. To see why, consider a capital improvement that is specific to task j. This improvement directly substitutes for labor in task j, with ν capturing the elasticity of substitution between capital and labor. At the same time, the productivity of the occupation has increased, and thus labor demand for its output; and χ is equal to the elasticity of demand for that occupation (Hicks 1932). The resulting impact on labor demand (and wage earnings) is a function of the sign of $\nu - \chi$, that is, whether the decline in labor demand due to improvements in labor-saving technology is greater than the increase in labor demand for the occupation as its productivity increases.

The second term in equation (10) captures the extent to which technological improvements are concentrated in specific tasks. This term depends on the degree to which technology improvements are concentrated in specific tasks,

$$C\left(\varepsilon\right) \equiv \sum\nolimits_{j\in J}\frac{\alpha\left(j\right)}{\sum\nolimits_{k\in J}\alpha\left(k\right)}\left(\varepsilon\left(j\right)-m\left(\varepsilon\right)\right)^2, \tag{13}$$

as well as

$$\eta_o = -\frac{s_k\beta\left(\nu-\psi\right)}{\left(1-\beta\right)+\beta\left(\nu s_k+\psi\left(1-s_k\right)\right)}, \tag{14}$$

which captures the impact of $\varepsilon(j)$ on the wage paid for task j $w(j)$ relative to the effect on $\varepsilon(j)$ on the wage paid on other tasks $j' \neq j$.

This concentration effect emerges due to two forces. First, in the model, workers can optimally respond to changes in labor productivity by reallocating time across tasks, with the scope for reallocation being inversely related to β. Second, log wages are a convex function of the vector of task-level wages, so Jensen's inequality implies that (appropriately weighted) mean preserving spreads in task prices log $w(j)$ increase occupation wages and labor demand. The more that productivity improvements are concentrated in a subset of tasks, both effects become larger quantitatively. The concentration effect is unambiguously positive, partially offsetting the (typically negative) direct effects associated with $m(\varepsilon)$.

The final two terms in equation (10) capture productivity spillovers. The third term captures increases in labor demand due to technology improvements. This term in turn depends on two economic forces: first, changes in α_I, for instance, due to the creation of new products; and second, changes in industry productivity Z_I, as the cost of production falls. The extent to which changes in industry productivity increase labor demand depends on

$$\eta_z \equiv \frac{\partial \log w(j)}{\partial \log Z_I} = \frac{\theta - \chi}{s_k \nu + s_l \chi + \zeta}. \tag{15}$$

As industry I becomes more productive, it will increase its labor demand for occupations. At the same time, however, whether the demand for the industry's output increases or decreases depends on θ. If the elasticity of substitution across industries θ is greater than the elasticity of substitution across occupations within a given firm χ, then an increase in industry productivity leads to an increase in the wage of each task. The final term in equation (10) captures aggregate labor demand and supply effects.

In the next section, we will construct direct empirical proxies for the first three terms in equation (10). Most importantly, our empirical design will include calendar year fixed effects; therefore, we cannot directly estimate the last term in equation (10). Given this "missing intercept" problem, our empirical regressions can only identify relative employment shifts across occupations and industries. This issue becomes even more salient in the pre-1910 sample, given that the census only included industry codes in 1910. This implies that, for the pre-1910 period, both the third and the fourth terms in equation (10) are absorbed.

II. Measurement

Our model in the previous section implies that the impact of a specific technology on labor demand for a specific occupation can be summarized by three objects: the average improvement in the labor-saving technology across the tasks performed by workers in that job, equation (11); the degree to which these improvements are concentrated in specific tasks, equation (13); and the impact of industry spillovers—the term involving Z_I

in equation (10). In this section, we describe how we construct empirical analogues of these objects in the data.

II.A. Data Sources

Our broad goal is to use the textual description of these innovations to relate these innovations to the tasks that workers do. Here, we briefly describe the sources of data for our empirical analysis. We relegate all details to the online appendix.

TECHNOLOGY IMPROVEMENTS We measure technology improvements using patents. We obtain the textual description of innovations of patents for the 1836–2024 period from Google Patents and PatentsView. Since the quality of the extracted text varies considerably over time, we use a modern LLM to obtain a concise description of the key new innovation from each patent document. To represent these patent summaries in numerical data, we use embeddings provided by OpenAI.[3] Online appendix B.1.1 and B.1.2 provide further details.

We next construct a coarse measure of which industries are particularly likely to see productivity improvements from technology *p*. In particular, as part of the LLM-generated summary in the previous section, we also query the LLM to provide a likely industry of use for each patent. Using the same OpenAI embeddings as above, we compute the cosine similarity between these industry-of-use descriptions from the LLM and the textual descriptions of the industries from the census (using the time-consistent IND1950 census industry scheme). We assign a patent as relevant to a given census industry if the census industry title is the most textually similar to the patent's industry-of-use description (based on cosine similarity of the embeddings generated for the patent industry of use and census industry title); we further assign the patent to the industry if the textual similarity is in the top 1 percent of all industry-patent similarity pairs for patents issued in that decade.[4] Online appendix B.1.3 provides further details.

As part of some of the descriptive analysis in the next section, we group technologies into broad categories over different time periods. To do so,

3. We use the OpenAI text-embedding-3-small model. Embeddings are geometric representations of the semantic meaning of text and have the property that the cosine similarity of two embedding vectors has a high value when the words have a similar meaning (see, e.g., Mikolov and others 2013).

4. For example, the LLM indicates that the industry of use for US patent 6,009,696 (harvester head for dried-on-the-vine raisins) is "most likely to be used in the agricultural industry, specifically in the production of vine crops such as raisins, olives, and other tree crops." Based off the textual similarity between this sentence and the census industry description, we match this patent to census industry 356, Agricultural Machinery and Tractors.

we estimate a *k*-means clustering algorithm on the document embeddings in order to separate patents into twenty groups over three distinct technology periods (1850–1920, 1920–1980, and 1980–2020). Then, we provide a large number of example summaries from each cluster to the OpenAI o3 model and ask it to provide concise labels summarizing each technology. See online appendix B.1.4 for further details.

TASKS A key challenge in our long-run empirical analysis is that the scope and function of occupations have changed over time. Given our use of the census data, an important constraint is the census definitions of occupations (which change over time). Unfortunately, a detailed description of what these occupations do is not available—the earliest vintage of the *Dictionary of Occupational Titles (DOT)* starts in 1939. To overcome this challenge, we employ a specialized LLM trained to understand and execute web search queries (GPT-4o search preview model), and direct the LLM to provide a list of tasks performed by each census occupation, in each decade from 1850 to 2010, in the style of occupation task descriptions in O*NET. The result of this query is a combination of 84,393 task-occupation-decade, spanning the years 1850–2010. On average, the search query returns approximately fourteen tasks per occupation in each decade. Online appendix table A.1 lists some examples. As before, we represent these tasks as numerical vectors using the same embeddings from OpenAI that we use for the technology descriptions in the previous section. Online appendix B.2.1 provides further details on the procedure.

These LLM-sourced task descriptions play a key role in our empirical analysis. To validate these descriptions, we compare their semantic meaning with either the *DOT* or O*NET counterparts whenever these are available. Specifically, for each occupation in the 1940, 1980, and 2010 decades, we compute the (cosine) similarity between the average task embedding of the LLM-sourced task descriptions and the average task embedding of its *DOT* or O*NET counterparts. We then compare the distribution of these similarity scores to a placebo distribution where we randomly compare the LLM tasks with a different occupation in the *DOT* or O*NET. As we see in online appendix figure A.1, the task descriptions we obtain from the LLM are fairly similar to their *DOT* or O*NET counterparts from the same period and are significantly different from the placebo distribution. Interestingly enough, the LLM tasks from the 2010 period are particularly close to the O*NET tasks, which suggests that these tasks are part of the LLM's search query. In brief, using this procedure allows us to measure worker tasks during periods where official sources are not easily available.

In addition, we use the same LLM (GPT-4o search preview) to obtain a classification of these tasks into three categories: manual, cognitive, and interpersonal. To help the LLM classify each task, we first manually classify a list of forty-one distinct occupational work activities from O*NET into a cognitive, manual, or interpersonal category. We then provide the LLM with these categorizations as context for how to categorize a task into one of the three categories. Approximately 45 percent of the task-occupation-year triplets are classified as manual, 34 percent as cognitive, and 21 percent as interpersonal tasks. See online appendix B.2.2 for details.

EMPLOYMENT We obtain employment counts using the IPUMS census extracts. We extract information on gender, age, occupation, industry, and labor force participation. We restrict the sample to census respondents age 15–75 who report that they are employed. We exclude members of the armed forces and occupation codes indicating a nonoccupational response. Following Katz and Margo (2014), we include both men and women in our analysis; however, we also investigate the robustness of our findings by restricting to men. We compute employment in a specific occupation (or industry, when available) using the census respondent weights. To compute employment growth across decades, we aggregate the decade-specific definition of census occupations into time-consistent classifications. Online appendix B.3 contains further details.

II.B. Measuring Technology Exposure

The next step is to estimate the direction of technological progress, specifically, the extent to which the tasks of a particular occupation are exposed to technological innovation in a given period.

TASK TECHNOLOGY EXPOSURE We measure the exposure of a task performed by a given occupation to a particular technology using textual similarity between the summary of patent documents and the LLM-sourced tasks described above. For each pair of patents p and tasks j in a particular decade T, we compute the cosine similarity between their OpenAI text embeddings. We consider a task j exposed to technology p if its cosine similarity exceeds a particular threshold: As our baseline, we consider the 95th percentile of the distribution.[5] Thus, we compute the probability that task j is exposed to technology improvements in the period from T to $T+H$ as

5. Time variation in our measure is confounded by changes in language over time. To address this, we demean patent-task similarity scores by year and compute percentile cutoffs separately for tasks in each census decade. Specifically, for a task j linked to a census occupation in decade T, we calculate the 95th percentile cutoff using the distribution of similarity scores across patents issued from twenty years before to twenty years after the start of decade T.

$$\text{Exposure}_{j,T}^{H} = \frac{1}{\left|P_{T,T+H}\right|} \sum\nolimits_{p \in P_{T,T+H}} \mathbf{1}\left(\text{similarity}_{p,j} > p95\right). \tag{16}$$

Equation (16) measures the direction of technological progress in a given decade: the average exposure of task j to technologies in patents issued between years T and $T+H$. Through the lens of the model, we can interpret equation (16) as being proportional to the degree of technology improvements $\varepsilon(j)$ that are specific to task j in a particular period.

OCCUPATION TECHNOLOGY EXPOSURE The next step involves extrapolating from task exposures to technology to the exposure of a particular occupation to technology. The model provides a useful guide—equations (11) and (13)—which states that an occupation's exposure to technology is a function of not only its average task exposure to technology, but also the extent to which this exposure is concentrated in particular tasks. We construct the direct analogues of equations (11) and (13) in the data as

$$\text{Mean Exposure}_{o,T}^{H} = \frac{1}{\left|J\left(o,T\right)\right|} \sum\nolimits_{j \in J\left(o,T\right)} \text{Exposure}_{j,T}^{H} \tag{17}$$

and

$$\text{Exposure Concentration}_{o,T}^{H} = \frac{1}{\left|J\left(o,T\right)\right|} \sum\nolimits_{j \in J\left(o,T\right)} \left(\text{Exposure}_{j,T}^{H} - \text{Mean Exposure}_{o,T}^{H}\right)^{2}. \tag{18}$$

Given our timing convention, equations (17) and (18) measure the flow of innovation across one or two decades.[6]

There are two important caveats that we should keep in mind. First, the measures of occupation exposure vary only at the occupation level; there is no variation across industries. The reason is that our assignments of

For example, for the 1940 Census tasks, the cutoff is based on patents issued between 1920 and 1959. Using alternative thresholds (e.g., the 90th or 99th percentile) yields nearly identical results. Because the cutoffs are decade-specific, low-frequency time series variation in our measure should be interpreted with caution. However, secular shifts in language are less likely to confound the cross-sectional dimension of our exposure measures.

6. The fact that occupation definitions change over time implies that we need to subtly modify equations (17) and (18) to map them into time-consistent census occupation codes. In the cases where multiple decade-specific occupation codes map into a single time-consistent occupation code (using the census OCC1950 scheme), we compute weighted averages of equations (17) and (18), using the share of individuals in each aggregated occupation.

patents to sectors using LLMs are a bit too coarse to accurately compute equations (17) and (18) at the sector-occupation level.

Second, our measure of task exposure, equation (16), which is used to construct equations (17) and (18), measures the direction of technological progress in a given decade. What is missing from these expressions is a measure of the *intensity* of technological progress. Given our use of patent data, it is quite challenging to identify shifts in the intensity of innovation that is directed to specific tasks—and potentially specific sectors. To overcome this limitation, we will also explore the relation between technology exposure and employment growth period by period.

INDUSTRY SPILLOVERS Given the long historical period we study, obtaining consistent measures of industry output or productivity is difficult. We therefore rely on patents, which we assign to industries using the LLM-based procedure described above. Because the industry matching is text-based and census industry descriptions are sparse, there are some industries that persistently get matched to a large share of patents (such as IND1950 code 399, Miscellaneous Manufacturing Industries), while others get matched to a small share. This creates large-level differences in the flow of patents matched to an industry. To account for this, we instead rely on growth rates in the number of patents assigned to a given industry. For twenty-year employment changes, our industry spillover measure is

$$\text{(19)} \qquad \text{Spill}_{I,T}^{20} = \log\left(\text{Matched Patents}_{T\to T+20,I}\right) - \log\left(\text{Matched Patents}_{T-20\to T,i}\right)$$

where Matched $\text{Patents}_{T\to T+20,i}$ is the number of patents issued between years T and $T + 20$, which are matched to industry I. The measure for the ten-year horizon $\text{Spill}_{I,T}^{10}$ is defined analogously. Accordingly, we map the broad changes in technology-related industry productivity implied by the model to the growth in the number of patents that are applicable to that industry.

A SHIFT-SHARE IV The direction of innovation is likely endogenous to the state of the labor market in a particular period. To address this concern we build a shift-share IV that builds on Acemoglu, Akcigit, and Kerr (2016). In particular, our shift-share identification strategy leverages the extent to which variation in the arrival of breakthrough technologies in an "upstream" technology class leads to the development of "downstream" (labor-saving) technologies in other tech classes. Our identification assumption is that the development of these upstream technologies is unrelated to the labor market conditions of the occupations exposed to the downstream technologies.

The construction of our shift-share instrument entails two steps. The first step involves predicting the arrival of the number of patents $N_{c,t}$ in a given tech class c at time t using a Poisson regression. Our predictor is constructed based on previous breakthrough innovations in other tech classes that occurred in the past:

$$\lambda_{c,t} = \sum_{\tau} \sum_{c' \neq c} \Omega_{c' \to c,t,\tau} \times I_{c',t-\tau}. \tag{20}$$

The predicted number of downstream patents $\lambda_{c,t}$ depends on two objects. First, $\Omega_{c' \to c,t,\tau}$ is a technology diffusion matrix constructed based on the textual similarity of patents: Its elements are the average similarity of patents in technology class c to patents in tech class c' for patents issued in tech class c at time t and tech class c' at time $t - \tau$; $\tau = 5, \ldots, 20$, which represents a diffusion lag for innovation to propagate from tech class c' to c. When constructing Ω, we set its diagonal to zero: We only use spillovers to tech class c from other tech classes c'. Second, $I_{c,t}$ is equal to the intensity of breakthrough patents in tech class c and year t, measured by the number of patents that are breakthroughs according to Kelly and others (2021) granted in year t and in tech class c relative to total patents in year t and tech class c. We estimate the Poisson regression at an annual frequency, allowing the estimated coefficients to vary by decade T, and then sum the predicted number of patents each decade. Overall, equation (20) is a strong predictor of subsequent patenting in tech class c, with an average t-statistic of 4.2 across the decade-by-decade coefficients. See online appendix C.4 for further details.

The second step involves estimating the likelihood that a patent p from tech class c is related to task j:

$$\alpha_{j,T} = \frac{1}{|P_{c,T-10}|} \sum_{p \in P_{c,T-1}} \mathbf{1}\left(\text{similarity}_{p,j} > p95\right). \tag{21}$$

When constructing the exposure shares, equation (21), we use all patents issued in the previous decade.

Putting the pieces together, our shift-share measure for the exposure probability, equation (16), of task j to technological innovation in period T is given by

$$Z^{H}_{j,T} = \sum_{c} \alpha_{j,T} \times \frac{\hat{N}^{H}_{c,T}}{|P_{T,T+H}|}. \tag{22}$$

Our instrument is equal to the sum across tech classes of the product of the predicted shifts in the direction of innovation, equation (21), times the predicted share of innovation in tech class c in decade T over the horizon H—the fitted values from the Poisson regression above divided by the total number of patents during this period. Using equation (22), we proceed to construct instruments for our mean (equation 17) and concentration (equation 18) measures, by replacing equation (16) with equation (22) in their definition.

Last, we instrument for industry-level technological spillovers as follows. The first step is to predict the level of patenting in the industry at the H-year horizon based on innovation in upstream technologies as

$$\text{Predicted Patents}_{T \to T+H,I} = \sum_c \Gamma_{I,c,T-10} \times \hat{N}^H_{c,T}, \tag{23}$$

where $\Gamma_{I,c,T-10}$ is the probability that a patent in time period T belonging to tech class c is relevant for industry I given our textual mapping of patents to industries. We compute the growth rate of equation (23) at the H-year horizon to instrument for our industry spillover measure, equation (19).

II.C. Examples

A key advantage of our measure is that it is available for essentially all years for which patents are available, and thus allows us to study very different technologies across long periods of time. Here, we first discuss some specific examples and then examine which major innovations are driving our exposure measures during different periods.

EXAMPLES FROM SPECIFIC INVENTIONS As an early example from the nineteenth century, consider the Bessemer process (US patents 16,082 and 49,055), the first method for mass-producing steel inexpensively. The occupations most exposed to these patents were in the iron and steel industries. Another example is the invention of the sewing machine. US patent 276,146, issued in 1883, is most closely linked to sewing machine operators, followed by carpet makers and thread makers. Other occupations had lower mean exposure overall but concentrated exposure in specific tasks, such as tailors and tailoresses, and shirt, cuff, and collar pattern makers. These examples illustrate how major breakthroughs in the Second Industrial Revolution directly transformed manual occupations in textiles and metalworking.

A particularly vivid case of skill displacement is the cylinder machine (US patent 814,612), a major innovation in the window glass industry during the late nineteenth century and included by Kelly and others (2021) in their list of breakthrough patents. Within a generation, the process of producing window glass shifted from handblown to mechanization, and entire artisanal branches of the industry disappeared. As Jerome (1934) documents, wages of blowers and gatherers fell by 40 percent, and multiple skilled trades were eliminated. We identify glassworks operatives as being among the most related occupations to this patent. Moving to the early twentieth century, Jerome (1934) describes the Barber-Colman warp-tying machine (patent 1,115,399), which could replace the work of about fifteen hand operators and be run by a single tender. In our methodology, the closest occupations include thread makers, sewing machine operators, lace makers, and carpet makers. Another example is the drawing-in machine (patent 1,364,091), which displaced five or six manual drawers-in with one operator and an assistant, again mapping closely to textile occupations. Other cases from this period further illustrate the link between particular patents and occupational exposure: the refrigerator (patent 1,276,612), linked to traders and dealers in ice; the modular combine harvester (patent 4,846,198), linked to farm laborers; the cotton picker spindle (patent 2,716,320), linked to textile weavers; the washing machine (patent 2,758,461), linked to laundresses; and the automatic telephone exchange systems (patents 1,146,583 and 1,439,723), linked to telephone operators.

Not all examples from this period represent labor-saving innovations. Some patents describe technologies that support cognitive tasks and could plausibly complement labor rather than substitute for it. For instance, patent 1,138,792 (a calculating machine) relates to the task, "prepare financial statements, such as balance sheets and profit and loss statements, by aggregating data from various ledgers and journals," performed by accountants and auditors. Likewise, patent 2,655,941 pertains to the task, "diagnose mechanical or hydraulic system failures to determine necessary repairs," performed by airplane mechanics and repairmen. These cases underscore that exposure in our measure does not necessarily imply substitution but can also capture technologies that enhance or complement human tasks.

Finally, in the second half of the twentieth century, labor-saving patents shifted toward white-collar and service work. For example, US patent 5,911,135 (system for managing financial accounts by a priority allocation of funds among accounts) is most closely linked to financial officers, accountants, and auditors. Patent 5,828,979 (automatic train control system and method) is related to locomotive operating occupations and railroad conductors and yardmasters. Patents 5,696,906 and 5,592,560, which cover e-commerce technologies, are most closely associated with billing clerks and sales occupations.

WHICH TECHNOLOGIES ARE DRIVING EXPOSURE? We next ask which types of technologies drive our exposure measures. To do so, we split the sample into three periods and assign patents within each to twenty broad technology groups, using the clustering and labeling procedure described in section II.A and online appendix B.1.4. Figure 1 reports, for each period, the share of overall exposure (dark grey bars) and the share of patents (light grey bars).

1850–1920. Panel A covers the Second Industrial Revolution. The largest source of exposure is woodworking machinery, reflecting the importance of agriculture and related industries early in the period. Textile and manufacturing machines and railroad technology are also key contributors, consistent with the replacement of artisanal production by mechanized factory production. By contrast, categories such as packaging and containers and household and leisure devices account for a sizable share of patents but little worker exposure.

1920–1980. Panel B shows a similar imbalance. Mechanization technologies continue to dominate worker task exposure, well beyond their patent share. Meanwhile, several important innovations—industrial chemical processes, chemical compositions and polymers, and electronics and communications systems—generate many patents but little direct exposure, suggesting they reshaped production without directly displacing tasks at the occupational level.

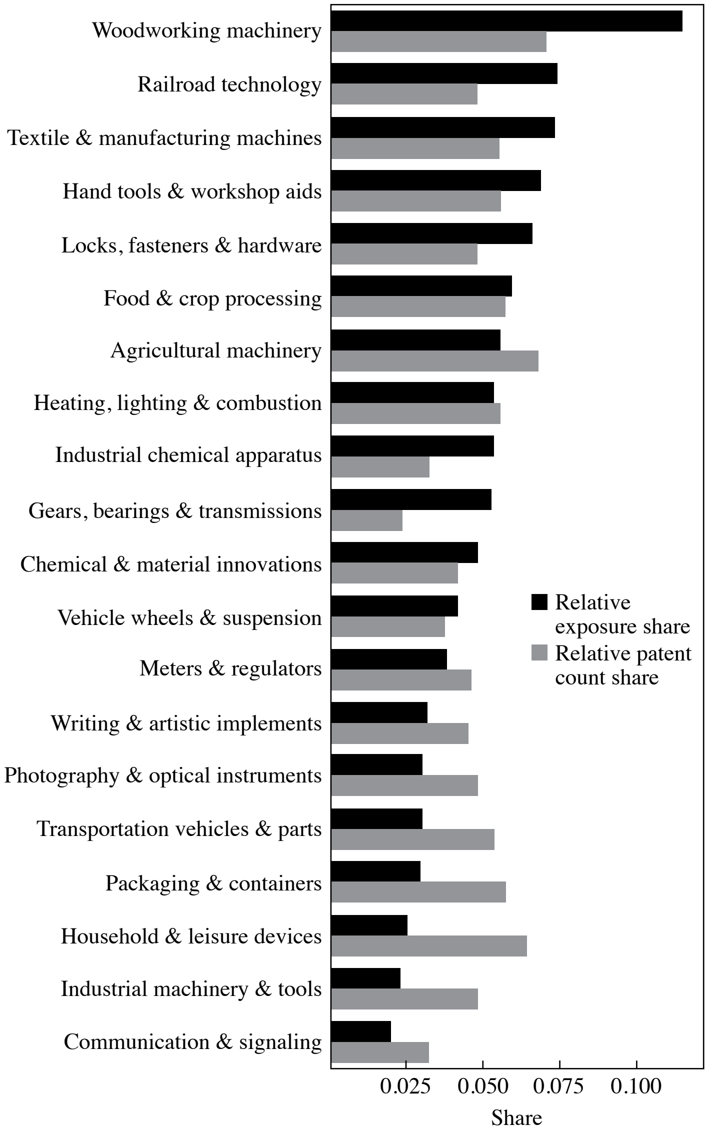
1980–2020. Panel C highlights the ICT era. Software, networking, and security systems emerge as the dominant source of worker task exposure, especially late in the sample. Earlier in the period, printing, paper, and copying contributed disproportionately. Across the period, ICT-related categories consistently drive exposure. Other breakthroughs, including advanced circuits and signal processing and pharmaceutical and biotech innovations, account for large patent volumes but a relatively small share of task exposure. A key difference from earlier eras is that exposure now increasingly comes from white-collar, cognitive technologies, in contrast to the manual, physical technologies that dominated earlier periods. We examine these occupational differences in more detail below.

III. Technology Exposure and Employment

Here, we discuss our main empirical findings regarding the impact of technology on the employment shares of different occupations. Section III.A presents our first set of results, which focus on identifying the direct effect of technology on labor demand, abstracting from productivity spillovers. Section III.B estimates the impact of productivity spillovers. Section III.C examines what our estimates imply for changes in the composition of

Figure 1. Composition of Overall Technology Exposure, by Technology Clusters

Panel A: 1850–1920



(*continued on next page*)

Figure 1. Composition of Overall Technology Exposure, by Technology Clusters (*Continued*)

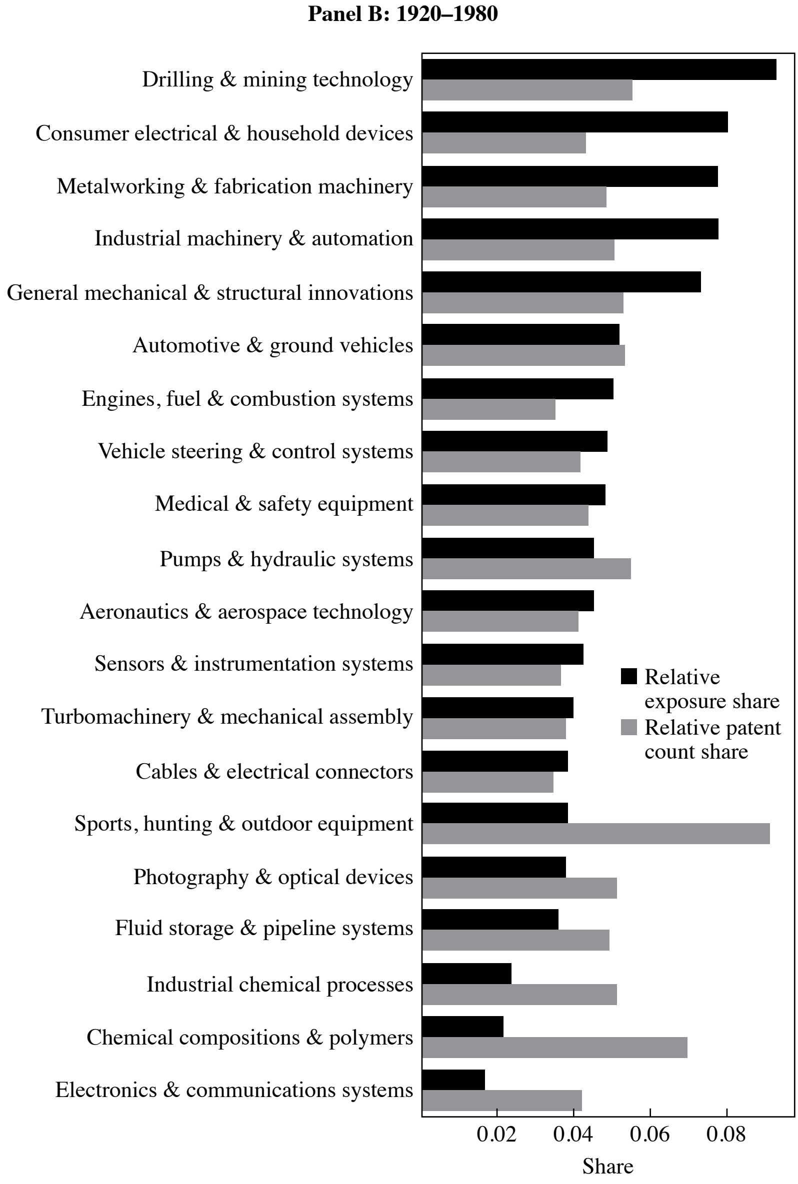
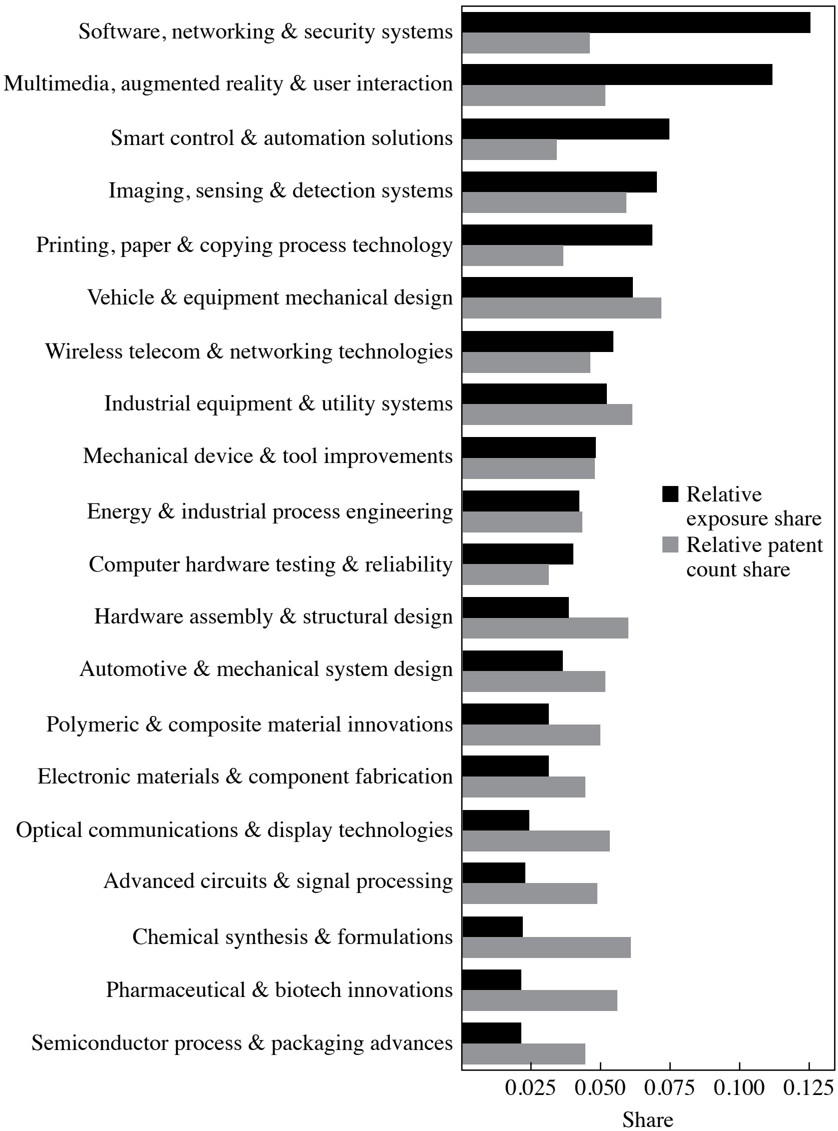


Figure 1. Composition of Overall Technology Exposure, by Technology Clusters (*Continued*)

Panel C: 1980–2020



Source: Authors' calculations.

Note: This figure plots the composition of technology exposure across twenty major technology clusters identified for the 1850–2020 period. Specifically, we use a *k*-means algorithm to separate technologies into twenty clusters, then use an LLM to generate short labels for each. The share of total occupational exposure and the fraction of patents associated with each cluster are shown in dark grey and light grey bars, respectively. Clusters are sorted vertically based on descending contributions to occupational exposure.

employment across occupations by education, average occupational wage level, and gender. Section III.D considers heterogeneity in the link between exposure and employment growth by different task types and across different age groups.

III.A. Technology Exposure and Occupation Labor Demand

We begin our analysis using the full sample of occupation-decade employment counts spanning from 1850 to 2020. Given the absence of consistent industry identifiers in the data, we are only able to identify the first two terms of equation (10). Thus, initially, we will abstract away from measuring the productivity effect of technology on labor demand.

BASELINE RESULTS We first examine the impact of the mean and concentrated technology exposure of occupation o to technology improvements in decade beginning year T over a horizon of H of ten or twenty years. In particular, we estimate

$$(24) \qquad \log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta \text{ Mean Exposure}^{H}_{o,T} + \gamma \text{ Exposure Concentration}^{H}_{o,T} + c\mathbf{\Gamma}_{o,t} + \varepsilon_{o,T}.$$

Equation (24) is the direct empirical analogue of equation (10) in the model. Here, $N_{o,T}$ denotes the employment of occupation o in census decade T. Given our timing convention, our ten-year specification relates the flow of innovation that occurred, for instance between 1850 and 1860, to the growth in employment over the same period. Depending on the specification, the vector of controls $\mathbf{\Gamma}$ includes calendar year effects, the share of occupation o in total employment in period T, and the occupation's employment growth over the last decade. Standard errors are clustered at the occupation level. Table 1 shows the estimated coefficients β and γ across different specifications.

First, we see that occupations with greater average exposure to technological improvements experience significant declines in employment growth relative to occupations with lower exposure. The coefficient β is consistently negative and statistically significant across all specifications, with magnitudes increasing at longer horizons. The ordinary least squares (OLS) and IV estimates are comparable. Economically, these effects are substantial: A one standard deviation rise in mean task exposure predicts an 11–13 percentage point employment decline over a decade—for comparison, the standard deviation of occupation employment growth is approximately 75 percent across a decade.

Table 1. Technology Exposure and Employment Growth, Direct Effects

	10-year horizon			*20-year horizon*		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Panel A: OLS						
Mean task exposure	−13.2***	−13.2***	−11.6***	−21.2***	−21.0***	−21.9***
	(1.92)	(1.93)	(2.26)	(3.33)	(3.34)	(3.56)
Concentration in task exposure	6.54***	6.36***	7.16***	8.90***	8.47**	12.8***
	(1.91)	(1.91)	(2.38)	(3.31)	(3.32)	(3.82)
Obs	3,212	3,212	2,452	3,166	3,166	2,410
R^2 (within)	0.019	0.026	0.046	0.024	0.035	0.068
Panel B: IV						
Mean task exposure	−13.5***	−13.5***	−11.4***	−23.7***	−23.8***	−21.6***
	(2.08)	(2.09)	(2.42)	(3.61)	(3.63)	(3.78)
Concentration in task exposure	6.38***	6.28***	6.31**	10.0***	9.82***	10.3**
	(2.12)	(2.12)	(2.62)	(3.65)	(3.66)	(4.19)
Obs	3,212	3,212	2,452	3,166	3,166	2,410
R^2 (within)						
F-stat (exposure)	6,606	6,624	5,554	4,282	4,288	3,663
F-stat (concentration)	2,222	2,229	1,865	1,487	1,491	1,302
Year FE	X	X	X	X	X	X
Employment share, lag		X	X		X	X
Employment share, lag growth			X			X

Source: Authors' calculations.

Note: The table above reports results from regressions of the form:

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta \text{ Mean Exposure}^{H}_{o,T} + \gamma \text{ Exposure Concentration}^{H}_{o,T} + c\boldsymbol{\Gamma}_{o,t} + \varepsilon_{o,T},$$

for decades T spanning from 1850 to 2000, excluding 1890. The variables of interest are Mean Exposure$^{H}_{o,T}$, technology mean exposure, and Exposure Concentration$^{H}_{o,T}$, technology exposure concentration (both normalized to unit standard deviation). Controls $\boldsymbol{\Gamma}_{o,t}$ include year fixed effects, lagged employment share $N_{o,T}$, and lagged employment growth $\log\left(\frac{N_{o,T}}{N_{o,T-10}}\right)$. The controls included in each regression specification are denoted by X. Coefficients are multiplied by one hundred. The top panel reports the estimated coefficients using OLS, while the bottom panel reports the IV estimates. Standard errors (in parentheses) are clustered by occupation. Columns 1–3 show regressions with $H = 10$, and columns 4–6 show regressions with $H = 20$. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

Second, controlling for mean exposure, increased concentration of technology exposure across tasks partially offsets these employment declines. That is, we see that the estimated coefficient γ is now consistently positive across different specifications and horizons, though it is somewhat less precisely estimated than the β coefficient. In terms of magnitudes, holding the degree of mean exposure constant, a one standard deviation increase in concentration is followed by approximately a 6–7 percentage point increase in employment over the next decade.

Third, the OLS and IV specifications are largely comparable across specifications. In general, there are two reasons to expect these two estimates to differ. First, there could be classical measurement error, which would tend to attenuate the OLS coefficients toward zero. Second, if labor-saving innovations are endogenously directed toward certain occupations that are expected to experience increased labor scarcity, then the OLS coefficient could be larger in magnitude. However, the fact that the IV coefficients are comparable in magnitude to the OLS estimates suggests that this type of selection is unlikely to be driving our estimates.

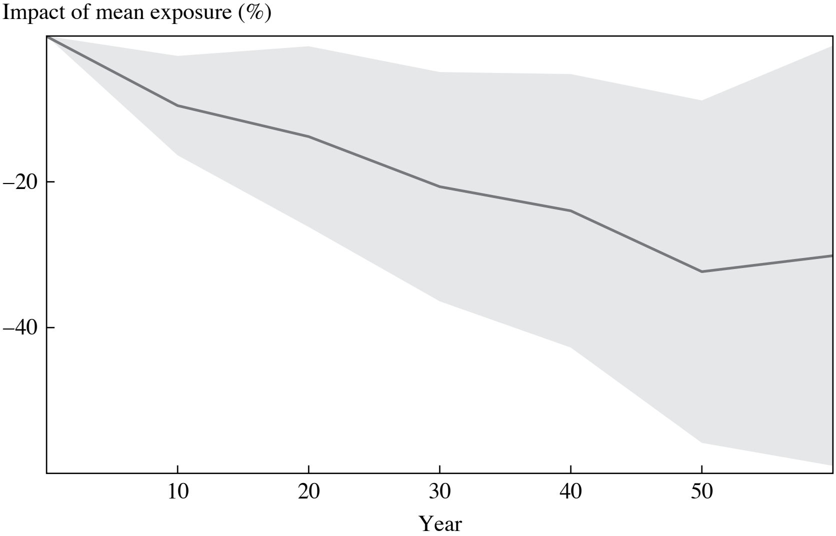
Fourth, interpreting the signs of these coefficients through the lens of the model, we see that the model indeed implies a positive estimate of γ. By contrast, a negative estimate of β strongly suggests that our measure of task exposure identifies technologies that substitute for labor in these tasks—that is, the elasticity of substitution between capital and labor ν is greater than the elasticity of substitution across tasks ψ.

Fifth, our baseline specification, equation (24), focuses on contemporaneous effects of technology exposure on employment growth. Even though we estimate long differences, delayed adoption may generate lagged responses. To account for this, we reestimate equation (24) using employment growth over longer horizons as the dependent variable, while holding exposure measures fixed at $H = 10$. Given the serial correlation of our mean and exposure measures (90 and 71 percent, respectively, at the ten-year horizon), we also include their lagged values to control for persistence in technological development. Figure 2 shows that estimated coefficients rise with the horizon but level off after about five decades, though standard errors widen given the limited sample. These results suggest that our baseline estimates likely understate the full extent of labor reallocation induced by technology.

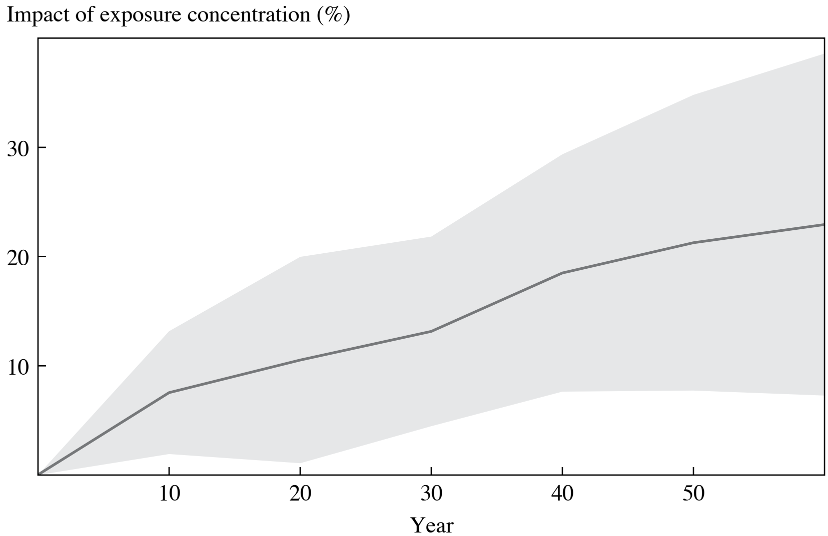
Last, recall from our discussion in section II.B, our mean and exposure measures only identify the direction of technological innovation in a particular decade. Absent a direct measure of the intensity of technological innovation that is consistently available over the entire period, we then

Figure 2. Technology Exposure and Employment Growth, Dynamic Effects

Panel A: Mean technology exposure



Panel B: Concentration of technology exposure



Source: Authors' calculations.

Note: This figure plots the coefficients from a long-horizon version of equation (24), in which the dependent variable is long-horizon growth in occupation employment. The specification uses our ten-year exposure measures as independent variables, and controls for one lag of the employment share and one lag of our exposure measures. We report the IV coefficients. Shaded bands represent 95 percent confidence intervals.

proceed to reestimate equation (24) every census decade. In online appendix figure A.2, we plot the estimated coefficients β and γ in each cross section that corresponds to column 5 in panel B of table 1.

Examining figure A.2, we see that the estimated coefficients β and γ have consistent signs and magnitudes, with the notable exception of during the 1880–1920 period in which they remain not really statistically different from zero. Online appendix table A.3 confirms this conclusion, by reestimating equation (24) across three subsamples: corresponding to technology improvements in the 1850–1920, 1910–1970, and 1960–2020 periods. Examining the table, we see that our estimates are statistically weaker during the 1850–1920 period. However, they are significant and in many cases have similar magnitudes in both the 1910–1970 and 1960–2020 subsamples.

Several explanations stand out for why our results are weaker in the 1880–1920 sample, which spans the Second Industrial Revolution. First, technological advances of this era may have been more complementary to labor than in later decades, particularly for skilled white-collar work (Goldin and Katz 1998). When $\nu \approx \chi$, the model predicts coefficients β and γ near zero, consistent with weak estimates. Second, technology may have affected labor demand beyond direct substitution in production. Improvements in transportation that expanded market access (Donaldson and Hornbeck 2016) likely boosted demand for manufactured goods. In the context of our model, these changes would manifest as shifts in $\alpha_I(o)$ in equation (2). Third, many forces beyond technology likely drove large shifts in both labor supply and demand during 1880–1910. Standardization of tasks and the division of labor expanded the pool of workers able to produce goods, while immigration between 1870 and 1914 supplied many laborers suited for factory work (O'Rourke, Taylor, and Williamson 1996; Abramitzky and Boustan 2017). Finally, measurement is noisier in this period. Census occupation codes were less consistent, and the harmonized series are coarse. Factory jobs often received distinct occupation codes from pre-mechanization artisan work. The 1910 Census marked a shift to a more standardized, function-based classification system, improving longitudinal comparability. For these reasons, in the rest of the paper we focus on the post-1910 period.

ROBUSTNESS CHECKS We perform several robustness checks to our main analysis. First, our baseline results pool male and female workers. One concern is that long-run changes in female labor force participation could affect our findings. Online appendix table A.5 shows this is not the case: Restricting the sample to men yields qualitatively and quantitatively similar results.

Second, we explore the sensitivity of our results if we were to obtain the tasks performed by each occupation using alternative LLMs. Our baseline uses GPT-4o search preview; we also generate tasks with GPT-4o and Llama 403B. Online appendix tables A.6–A.8 show that these models perform somewhat worse during 1850–1920 but lead to comparable results in 1910–1970 and 1960–2020.

Third, we reestimate our baseline regression weighting occupations by their employment shares, right-winsorized at the top 5 percent to avoid domination by very large occupations. Online appendix table A.4 shows the results are similar to the unweighted baseline.

Fourth, our empirical measures of the mean and concentration in task exposure are fairly highly correlated in the cross section—occupations with high degrees of technology tend to have concentrated task exposures. This raises the question of how the estimates of β change if we just focus on the first-order effects. As we see in online appendix table A.2, including only the mean exposure measure leads to smaller estimates of β: A one standard deviation increase in mean exposure is now followed by approximately a 6–7 percentage point decline in employment.

Fifth, we test whether our measures simply capture broad secular shifts in labor demand, such as the decline of manual work. As we show later, manual tasks are more exposed to technology than other task types, so falling demand for manual labor could confound our results. We therefore explore specifications that add controls for each occupation's initial share of manual and cognitive tasks in the 1910–2020 sample. Online appendix table A.9 shows that coefficients on mean exposure shrink modestly and become less precise, but remain significantly different from zero. Coefficients on concentration, by contrast, fall sharply and lose precision, suggesting this measure partly reflects exposure differences already embedded in broad task categories.

Finally, we address the possibility that our estimates reflect industry-level trends such as the secular decline of manufacturing. We reestimate equation (24) at the occupation-industry-decade level, restricting attention to the post-1910 period when industry codes are available. Online appendix table A.10 shows that including industry trends produces somewhat smaller coefficients but otherwise similar results.

III.B. Estimating the Impact of Technology Spillovers on Labor Demand

We next turn our attention to estimating the impact of productivity spillovers on labor demand. Doing so requires a modification to our empirical analysis: We now estimate our coefficients at the occupation-industry-decade level. Since industry classifications are only reliably available in the census post-1910, we restrict attention to this period. In particular, we estimate

$$(25)\quad \log\left(\frac{N_{o,I,T+H}}{N_{o,I,T}}\right) = \beta \text{ Mean Exposure}^{H}_{o,T} + \gamma \text{ Exposure Concentration}^{H}_{o,T} + \delta \text{ Spill}^{H}_{I,T} + c\mathbf{\Gamma}_{o,I,t} + \varepsilon_{o,T},$$

where $\text{Spill}^{H}_{I,T}$ is our measure of productivity spillovers in industry I in year T and horizon H. Now, $N_{o,I,T}$ denotes the employment of occupation o in census decade T in industry I, where industries refer to the census industry classifications (IND1950). As in our baseline results, the vector of controls $\mathbf{\Gamma}$ includes calendar year effects and the share of occupation o in total employment in period T. In addition, we include sector level (or sector interacted with decade) fixed effects. Sectors correspond to sixteen broad industry categories derived from the IND1950 classification. Interacting these sector-level fixed effects with decade fixed effects allows us to partial out broad trends across sectors (for example, services versus manufacturing) while still at the same time identifying productivity spillovers across the more granular industries within each sector.

Table 2 presents our estimates of β, γ, and δ. Examining the table, we note that the coefficient δ of employment growth is positively and significantly related to our productivity spillover measure. Focusing on panel A, which reports the OLS estimates, a one standard deviation increase in our spillover measure is associated with approximately an 8–10 percentage point increase in employment growth over the next decade, and a 17–20 percentage point increase over the next two decades. Examining the IV estimates in panel B, we note that our instrument for spillovers is quite weak in the ten-year period, hence we view these estimates as not reliable. The instrument is somewhat stronger at a horizon of twenty years, in which case the IV estimates are close to the OLS estimates. Last, the coefficients β and γ are both strongly statistically significant across specifications, though their magnitudes are a bit smaller than our baseline estimates.

The estimated coefficient δ is related to the third term in equation (15) capturing industry spillovers. This term is a function of the extent to which new technologies increase the number of new products at the industry level—the α_I term—and also depends on the relative value of the elasticity of substitution across industries θ compared to the elasticity of substitution across occupations within an industry χ, together with the magnitude of productivity improvements at the industry level.

Table 2. Technology Exposure and Employment Growth, Industry Controls and Innovation Spillovers

	10-year horizon				*20-year horizon*			
Employment growth (%)	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Panel A: OLS								
Mean task exposure	−4.22***	−4.20***	−7.96***	−8.02***	−7.23***	−7.19***	−15.0***	−15.0***
	(1.23)	(1.22)	(2.05)	(2.05)	(2.24)	(2.24)	(3.35)	(3.36)
Concentration in task exposure			4.53**	4.64**			9.51***	9.56***
			(1.86)	(1.86)			(2.76)	(2.77)
Industry spillover	8.16***	10.6***	8.17***	10.6***	17.3***	20.4***	17.3***	20.4***
	(2.45)	(1.78)	(2.45)	(1.78)	(3.37)	(3.39)	(3.37)	(3.39)
N	137,192	137,192	137,192	137,192	126,463	126,463	126,463	126,463
R^2 (within)	0.005	0.006	0.006	0.007	0.013	0.013	0.014	0.015

(*continued on next page*)

Table 2. Technology Exposure and Employment Growth, Industry Controls and Innovation Spillovers (*Continued*)

	10-year horizon				*20-year horizon*			
Employment growth (%)	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Panel B: IV								
Mean task exposure	−4.82***	−4.96***	−8.08***	−8.25***	−9.55***	−9.47***	−16.2***	−16.2***
	(1.22)	(1.21)	(2.06)	(2.06)	(2.15)	(2.15)	(3.21)	(3.23)
Concentration in task exposure			4.06**	4.10**			8.44***	8.59***
			(1.88)	(1.90)			(2.83)	(2.84)
Industry spillover	44.5**	43.3**	44.5**	43.3**	15.3*	19.1**	15.3*	19.1**
	(20.39)	(17.18)	(20.39)	(17.18)	(7.94)	(7.40)	(7.94)	(7.40)
N	135,637	135,637	135,637	135,637	125,956	125,956	125,956	125,956
R^2 (within)								
F-stat (exposure)	4,151	4,168	2,892	2,903	2,295	2,370	1,618	1,676
F-stat (concentration)			1,100	5,384			760	796
F-stat (spillover)	4	3	7	29	24	21	16	40
Year FE	X		X		X		X	
Sector FE	X		X		X		X	
Year × sector FE		X		X		X		X
Employment share, lag	X	X	X	X	X	X	X	X

Source: Authors' calculations.

Note: The table above reports results from regressions of the form:

$$\log\left(\frac{N_{o,I,T+H}}{N_{o,I,T}}\right) = \beta \text{ Mean Exposure}^{H}_{o,T} + \gamma \text{ Exposure Concentration}^{H}_{o,T} + \delta \text{ Spill}_{I,T} + c\boldsymbol{\Gamma}_{o,I,t} + \varepsilon_{o,I,T},$$

for decades T spanning 1910–2000. Controls $\boldsymbol{\Gamma}_{o,I,t}$ include year fixed effects, broad sector fixed effects (or year × sector fixed effects), and lagged employment share. Sectors that correspond to sixteen broad industry categories are derived from the IND1950 classification: Agriculture, Forestry, and Fishing; Mining; Construction; Durable Goods Manufacturing; Nondurable Goods Manufacturing; Transportation; Telecommunications; Utilities and Sanitary Services; Wholesale Trade; Retail Trade; Finance, Insurance, and Real Estate; Business and Repair Services; Personal Services; Professional and Related Services; and Public Administration. Coefficients correspond to a one standard deviation change in the independent variable and are multiplied by one hundred. Standard errors (in parentheses) are clustered by occupation and industry. Panel A reports the estimated coefficients using OLS, while panel B reports the IV estimates. Columns 1, 2, 5, and 6 show estimates excluding exposure concentration measure, while the others include it. Columns 1–4 show regressions with $H = 10$, while the others show regressions with $H = 20$. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

Last, we explore the extent to which these estimates are consistent across different subsamples. We find that our estimates are broadly comparable across the 1910–1970 and 1960–2020 periods, as we can see in online appendix tables A.12 and A.13. Further restricting the sample to only male workers leads to very similar results, as we see in online appendix table A.11.

III.C. Implications for Shifts in Composition of Labor Demand

So far, our estimates show how technology shifts relative labor demand across occupations. We now ask what these estimates imply for the changing composition of labor across different types of occupations. To do this, we reestimate equation (25) decade by decade, focusing on a twenty-year horizon and plotting the resulting OLS estimates corresponding to column 8 of table 2.[7] This approach lets us account for the fact that the mean and exposure measures capture the direction of innovation but not necessarily its intensity. Online appendix figure A.3 shows the estimates over time. The only real outlier is 1910–1940, when coefficients are especially large—consistent with Field (2003), who argues this was the most technologically intensive period of the twentieth century. This decade-by-decade approach highlights not just the direction of innovation, but also the intensity of technology's impact on employment over time.

Next, we use the time-varying estimates of β, γ, and δ from equation (25) to construct predicted employment share growth at the occupation level:

$$\begin{aligned}(26)\quad \text{Tech-Predicted Growth}_{o,I,T\rightarrow T+20} = {}& \hat{\beta}\ \text{Mean Exposure}^{H}_{o,T} \\ &+ \hat{\gamma}\ \text{Exposure Concentration}^{H}_{o,T} \\ &+ \hat{\delta}\ \text{Spill}^{H}_{I,T}.\end{aligned}$$

We compare this to actual employment share growth $\log(N_{o,I,T+20}) - \log(N_{o,I,T})$.

We then aggregate these occupation-level growth rates at a broader occupation group level using the start-of-period employment share of each occupation in the category. We group occupations in three ways: based on average educational attainment on each occupation; based on the female share of employment; and ten broad occupational groups sorted by average

7. The IV estimates for the mean and concentration exposure measures are close to the OLS results. The IV for the spillover measure is somewhat underpowered, especially in decade-by-decade estimation, so those coefficients are less stable.

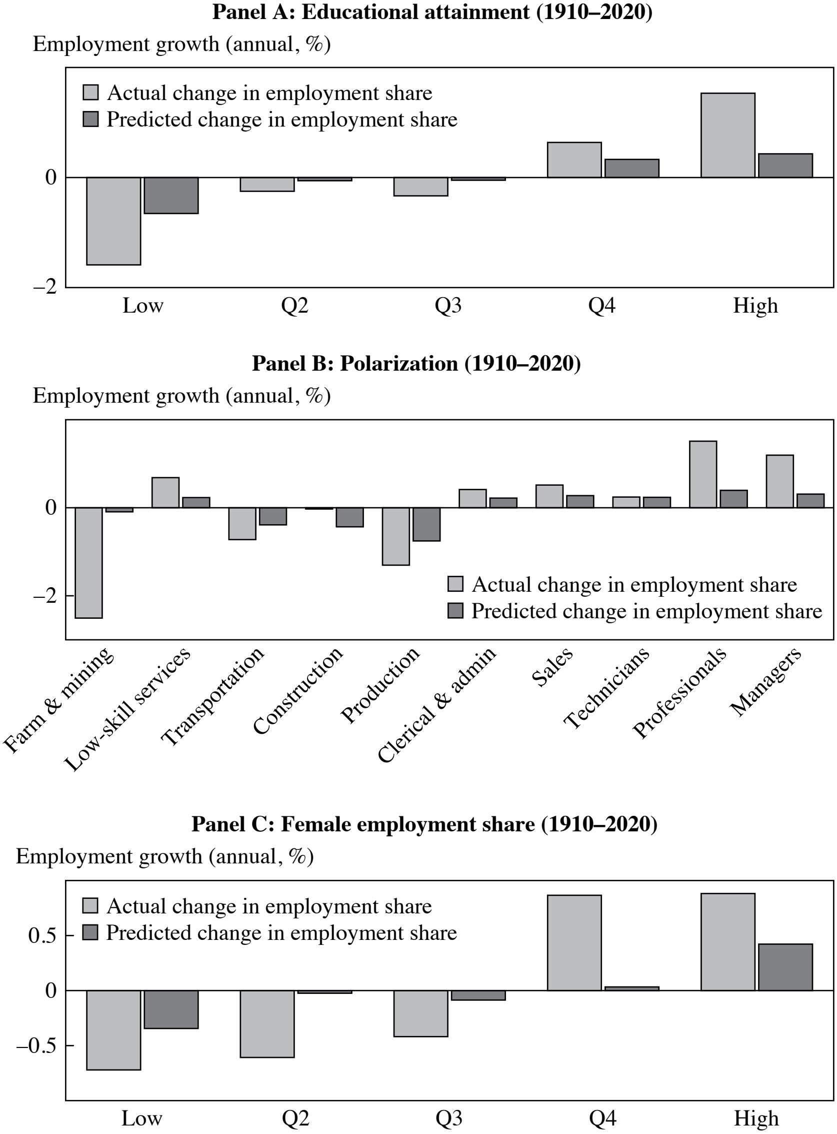
earnings, following Autor and others (2024). To summarize across decades, we pool within group *g* and report averages of predicted and actual annual growth in employment shares. In all cases, we demean by subtracting the across-group average so that the pooled values sum to zero in each decade. This exercise allows us to assess how well technology explains long-run differences in employment growth by education, gender, and broad occupation group.

BY WORKER SKILL (EDUCATION) The mainstream view is that twentieth-century technology was skill-biased: It raised demand for high-skill workers relative to low-skill ones (Goldin and Katz 2008; Katz and Margo 2014; Card and Lemieux 2001). Berman, Bound, and Machin (1998) and Krusell and others (2000) formalize this pattern as skill-biased technological change. We test this idea by estimating predicted employment growth by education. We measure education requirements within occupations using the IPUMS EDSCOR50 prior to 1980 and EDSCOR90 afterward—the share of workers with at least one year of college within each census occupation code. Panel A of figure 3 plots the mean predicted employment growth by quintiles of education, for the full 1910–2020 sample. Online appendix figure A.4 breaks this into subperiods. Each panel also shows the realized employment changes.

Panel A of Figure 3 shows that, over the last century, technology raised demand for high-education occupations relative to low-education ones. The effect is large: Our estimates imply that technological progress increased the relative demand for highly educated occupations by approximately 1.1 percentage points per year compared to the demand for occupations with lower levels of education. But the difference shrinks over time, falling to 0.9 percentage points in 1980–2020 as we see in panel A of table 3. A likely reason is the ICT revolution: As we will show in section III.D, we find that ICT increasingly substitutes for cognitive skills. Since more highly educated occupations also tend to do more cognitive tasks, this increased degree of substitutability likely led to relatively lower labor demand for these types of workers.

Panel A of table 3 shows that over the full sample, the gap in realized employment growth is approximately 3 percentage points per year between the highest and the lowest occupation education score groups; our estimates of exposure to technological progress can account for a third of that gap over this period. Moreover, the top quintile of occupations has gained in employment share relative to the bottom quintile across all subsamples; depending on the time period, technology explains between 27 and 47 percent of the change.

Figure 3. Technology Exposure and Shifts in Labor Demand Across Occupations



Source: Authors' calculations.

Note: This figure plots actual and technology-predicted average growth rates in employment shares based off estimates of equation (26), and also by occupational educational rank, following the procedure described in section III.C. In panel A, we sort occupations into yearly employment-weighted quintiles based off educational attainment (IPUMs variable EDSCOR50 for years before 1980 and EDSCOR 90 for 1980 onward). In panel B, we sort occupations into broad time-consistent groups following Autor and others (2024); occupation groups are sorted from left to right based off their average wages. In panel C, we group occupations based on the share of female workers in the occupation.

Table 3. The Role of Technology in Accounting for Observed Shifts in Labor Composition

	Full sample	*Subperiods*		
	1910–2020	*1910–1960*	*1950–1990*	*1980–2020*
Panel A: Occupation education level				
Employment growth (annual, %)				
Realized: high – low	3.1	4.0	3.1	2.0
Predicted: high – low	1.1	1.1	1.2	0.9
% of gap explained	34.8	27.5	39.4	47.4
Panel B: Occupation wages				
Employment growth (annual, %)				
Realized: low-skill – middle-skill	1.7	1.0	1.3	2.5
Predicted: low-skill – middle-skill	0.8	0.4	0.7	1.6
% of gap explained	49.6	35.1	58.0	63.2
Employment growth (annual, %)				
Realized: high-skill – middle-skill	2.1	1.7	2.5	2.1
Predicted: high-skill – middle-skill	0.9	0.5	1.1	1.4
% of gap explained	43.3	27.2	44.4	65.8
Panel C: Occupation female employment share				
Employment growth (annual, %)				
Realized: high – low	1.6	1.6	2.4	0.8
Predicted: high – low	0.8	0.5	0.7	1.2
% of gap explained	47.8	31.1	29.9	145.2

Source: Authors' calculations.

Note: In panel A of this table, we compare the average annualized actual and technology-predicted relative employment growth of occupations in the top employment-weighted quintile of educational attainment to the bottom quintile. In panel B, we compare the average actual and technology-predicted growth of middle-skill occupations relative to either low-skill services or high-skill occupations, using an occupational categorization inspired by Autor and others (2024). Finally, in panel C, we compare the relative actual and technology-predicted employment growth of occupations in the top employment-weighted quintile to those in the bottom quintile based off the share of workers who are female. Predicted employment growth comes from estimates of online appendix figure A.3. See section III.C for more details.

Our findings relate to Goldin and Katz (2008), who emphasize that the college wage premium fluctuated: high early in the century, compressed mid-century, then soaring after 1980. They attribute this to supply as well as demand. Our estimates show that technology consistently raised relative demand for educated occupations. But realized employment growth of educated occupations was largest in 1910–1960—precisely when our measures explain the least. This is consistent with supply-side factors (such as government subsidization of education post-WWII) causing a temporary increase in the growth of educated labor supply that outpaced pure technological considerations, leading to a temporarily deflated college premium.

BY OCCUPATION WAGES We next examine the degree to which our estimates can account for the observed degree of job polarization in the twentieth century. Following Autor and others (2024), we group occupations into ten broad categories, ordered left to right from low to high wage in figure 3. Unlike the twelve-group scheme in Autor and others (2024), we collapse the three small service categories into one low-skill services group, given the small size of health services early in the panel. Panel B of figure 3 covers the 1910–2020 period; panel B of table 3 and online appendix figure A.5 split the sample.

Examining the bars in the middle, and consistent with the literature on polarization (Autor, Katz, and Kearney 2006; Goos, Manning, and Salomons 2014; Autor and Dorn 2013), we find a decline in employment shares in low- and middle-skill occupations, like production, construction, and transportation occupations, across the whole time period. Technology-predicted growth in employment shares tracks these declines closely. At the same time, low-skill services on the left and high-skill occupations—professionals, managers, and, to a lesser extent, sales and technicians—grow in relative terms. Again, technology exposure lines up with these shifts.

These patterns are consistent across subperiods. Production declines in relative terms in every subsample, transportation in most. Our measures capture these patterns well. The post-1980 era shows the sharpest polarization, as highlighted in the literature, but we see that the process starts earlier. Consistent with Bárány and Siegel (2018), we find evidence of polarization in the post-1950 period, with a shift toward low-skill services and high-skill professionals (panel C of online appendix figure A.5). Further we see in panel B of figure A.5 that this pattern can even be traced earlier to the 1910–1960 period, suggesting that technology-related polarization has been an ongoing phenomenon for a longer period of time than previously recognized, even for up to a century.

To dig deeper, we focus on middle-skill, manual jobs—production, transportation, and construction—compared to low-skill services and to high-skill occupations (technicians, professionals, and managers). Including clerical and administrative jobs with middle-skill or sales with high-skill leads to similar conclusions. Table 3, panel B, reports relative changes. Low-skill services grew 1.7 percent per year relative to middle-skill jobs, while technology predicts 0.8 percent. Technology thus explains about half of the shift. Across subperiods, it explains between 35 and 63 percent of the change, depending on the time period in question. In the bottom three rows of panel B, table 3, we examine the other half of the polarization phenomenon: the shift toward high-skill occupations and away from these middle-skill occupations. From 1910 to 2020, high-skill jobs gained 2.1 percent per year relative to middle-skill jobs; technology explains 43 percent. By subperiod, the share explained ranges from 27 to 66 percent. In both cases, the contribution of technology is larger in the post-1980 era.

BY OCCUPATION FEMALE SHARE Female labor force participation has shifted dramatically over the past 150 years. As Ngai, Olivetti, and Petrongolo (2024) document, women's employment followed a *U*-shaped path: It fell as the economy moved out of farming, then rose again as employment shifted from manufacturing into services. Over the twentieth century, the expansion of clerical, office, and later, service jobs, was central to rising female employment and hours (Goldin 2006).

Our results show that technological progress over the past century has consistently favored white-collar service jobs over blue-collar production. A natural next question is whether these same forces can explain shifts in labor demand across occupations that differ by gender composition. We therefore compare predicted employment growth from our technology measures across occupations grouped by their initial female share. Panel C of figure 3 covers 1910–2020; online appendix figure A.6 examines subperiods separately.

Two patterns stand out. First, occupations with higher initial female shares grew substantially relative to male-dominated ones. Second, technology exposure explains much of this growth. Panel C of table 3 shows that, from 1910 to 2020, female-intensive occupations gained about 1.6 percentage points per year relative to male-intensive occupations. Our measures explain about half of that—0.8 percentage points per year. Looking by subperiod, technology accounts for roughly one-third of the shift in both 1910–1960 and 1950–1990, and somewhat more than the observed shift during 1980–2020.

In sum, we find evidence that technology pulled women into the labor market, and it did so long before 1980. Our estimates show that demand shifted steadily toward female-intensive occupations as technology favored services over production. The mechanism we emphasize operates through labor demand, not just supply. Of course, technological change also likely boosted female labor supply by substituting for home production (see Greenwood, Seshadri, and Yorukoglu 2005; Albanesi and Olivetti 2016; Bose, Jain, and Walker 2022). But the demand story is clear: Occupations with more women grew faster, and technology explains a large share of that growth. This pattern links directly to our polarization results. As Ngai and Petrongolo (2017) argue, the service economy underpins rising female labor force participation. Our measures provide the technology-based mechanism for that argument, and they show the shift started as early as 1910–1960, not only after 1980.

III.D. Heterogeneity

We next turn to heterogeneity, asking how exposure differs across tasks and across age groups.

HETEROGENEITY BY TASK TYPES Our analysis so far has implicitly assumed that the elasticity of substitution between capital and labor is similar across different types of tasks. This may not be a particularly good assumption: Technologies that are semantically similar to the description of interpersonal tasks are unlikely to have the same degree of substitution as technologies that are similar to manual tasks. Further, there are probably reasons to think that these elasticities of substitution change over time.

To explore this idea further, we next decompose our mean exposure measure into three components, that is, the ones arising from the technology exposure of manual M, cognitive C, and interpersonal I tasks. Specifically, the mean exposure measure is equal to the sum of the exposure of the manual, cognitive, and interpersonal mean task exposure, multiplied by their respective task shares,

$$\text{Mean Exposure}^{H}_{o,T} = \sum\nolimits_{\tau \in \{M,C,I\}} \text{Task Share}^{\tau}_{o,T} \times \text{Mean Exposure}^{\tau,H}_{o,T}, \tag{27}$$

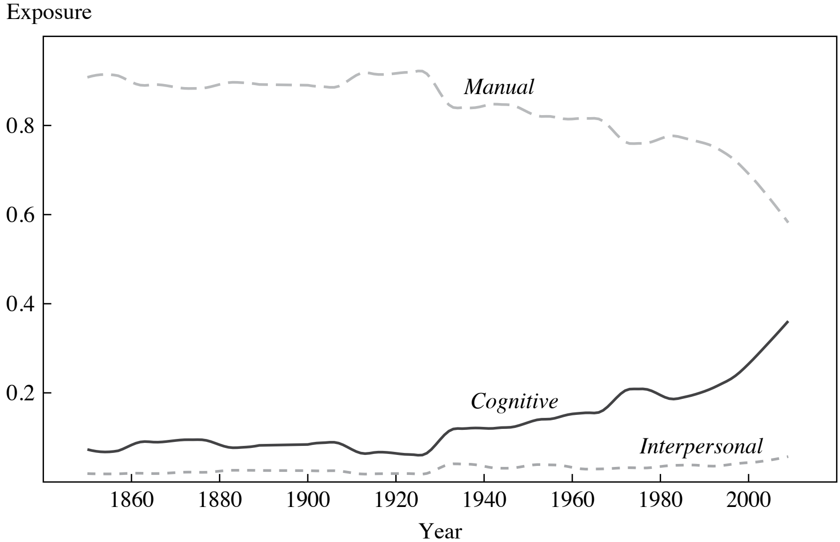
where each task-type $\tau\{M, C, I\}$ mean exposure is equal to

$$\text{Mean Exposure}^{\tau,H}_{o,T} = \frac{1}{\left|J(o, T, \tau)\right|} \sum\nolimits_{j \in J(o,T,\tau)} \text{Exposure}^{H}_{j,T}, \tag{28}$$

and $J(o, T, \tau)$ denotes the tasks of type τ performed by occupation o in decade T. These measures are motivated by the model of Kogan and others (2023), who show that they emerge as key sufficient statistics for worker exposure in a model with CES production and tasks that differ in terms of their elasticities of substitution between capital and labor (ν in our model).

Figure 4 shows how the composition of which workers are most exposed has shifted across different task categories over the sample period. For each year, we measure how technological innovation is distributed across different types of work by calculating the share of high-similarity patent-task links that involve manual, cognitive, or interpersonal tasks. Examining the figure, a few points stand out. First, throughout the entire sample, occupations emphasizing manual tasks are overall much more exposed to

Figure 4. Composition of Overall Technology Exposure, by Task Type



Source: Authors' calculations.

Note: This figure shows how the share of technology exposure varies across task types—manual, cognitive, and interpersonal—over time. For each year t, we compute the fraction of all high-similarity patent-task links that involve tasks of a given type τ. In particular, c_τ is the share of all valid patent-task matches exceeding the 95th percentile similarity threshold, which correspond to type τ tasks. The figure plots these shares, smoothed using a LOWESS filter (bandwidth = 0.1), for all years 1850–2010.

automation than occupations emphasizing cognitive or interpersonal tasks. This is consistent with the traditional view that automation primarily affects workers in manufacturing industries (blue-collar workers). However, the ICT revolution has likely expanded the set of occupations that are affected to white-collar workers. Indeed, as we see in the figure, there is an increasing trend in the exposure of occupations emphasizing cognitive tasks since the mid-1980s. Online appendix figure A.7 shows the breakdown of manual, cognitive, and interpersonal exposure across major clusters. ICT-related technologies, with software appearing at the top of the list, emerge among the set of technologies with the highest exposure to cognitive and interpersonal tasks.

Armed with this decomposition, we then reestimate equation (24) and replace our mean exposure measure with its individual components. Table 4 presents the estimated coefficients. Panel A focuses on the post-1910 period, while panels B and C report results separately for the 1910–1970 and 1960–2020 periods. Given that, in the context of the model, decomposing

Table 4. Technology Exposure and Employment Growth, Subsamples, Heterogeneous Effects by Task Exposure

	OLS				*IV*			
	10-year horizon		*20-year horizon*		*10-year horizon*		*20-year horizon*	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Panel A: Full sample (1910–2020)								
Mean task exposure, manual	−2.15***	−2.92***	−4.70***	−6.33***	−2.29***	−2.85***	−5.00***	−6.30***
	(0.26)	(0.46)	(0.51)	(0.82)	(0.26)	(0.49)	(0.51)	(0.86)
Mean task exposure, cognitive	0.74	−0.59	1.42	−1.50	0.58	−0.37	0.85	−1.47
	(0.69)	(1.01)	(1.38)	(1.81)	(0.67)	(1.06)	(1.30)	(1.91)
Mean task exposure, interpersonal	−0.41	−1.20	−2.53	−4.20	−0.59	−1.17	−2.86	−4.03
	(2.22)	(2.23)	(4.68)	(4.80)	(2.26)	(2.26)	(4.59)	(4.71)
Concentration in task exposure		10.3**		22.5**		7.57		18.2*
		(5.25)		(9.09)		(5.67)		(9.89)
Obs	2,436	2,436	2,375	2,375	2,436	2,436	2,375	2,375
Panel B: 1910–1970								
Mean task exposure, manual	−1.77***	−2.46***	−4.29***	−6.07***	−1.94***	−2.55***	−4.54***	−5.84***
	(0.40)	(0.74)	(0.75)	(1.18)	(0.41)	(0.83)	(0.77)	(1.30)
Mean task exposure, cognitive	5.29***	3.97**	7.32***	3.70	5.08***	3.94*	6.64***	4.06
	(1.50)	(2.00)	(2.64)	(3.03)	(1.34)	(2.01)	(2.29)	(2.95)
Mean task exposure, interpersonal	−1.07	−1.48	−6.62	−7.44	−1.45	−1.87	−6.09	−6.78
	(2.75)	(2.77)	(7.33)	(7.15)	(2.82)	(2.88)	(7.05)	(6.95)
Concentration in task exposure		9.14		23.8*		8.28		17.8
		(9.16)		(13.12)		(10.83)		(15.76)
Obs	1,116	1,116	1,153	1,153	1,116	1,116	1,153	1,153

(*continued on next page*)

Table 4. Technology Exposure and Employment Growth, Subsamples, Heterogeneous Effects by Task Exposure (*Continued*)

	OLS				*IV*			
	10-year horizon		*20-year horizon*		*10-year horizon*		*20-year horizon*	
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Panel C: 1960–2020								
Mean task exposure, manual	−2.57***	−3.09***	−5.07***	−5.85***	−2.78***	−3.26***	−5.36***	−6.36***
	(0.33)	(0.54)	(0.63)	(0.98)	(0.33)	(0.57)	(0.62)	(1.07)
Mean task exposure, cognitive	−1.31**	−2.14**	−2.11*	−3.35	−1.52**	−2.31**	−2.49*	−4.13*
	(0.63)	(1.06)	(1.27)	(2.20)	(0.65)	(1.15)	(1.29)	(2.42)
Mean task exposure, interpersonal	0.69	−0.33	4.29	2.68	0.12	−0.62	2.23	0.69
	(4.17)	(4.00)	(6.88)	(7.21)	(4.24)	(4.08)	(7.29)	(7.73)
Concentration in task exposure		7.44		10.9		6.93		14.0
		(5.38)		(10.23)		(5.91)		(11.89)
Obs	1,320	1,320	1,222	1,222	1,320	1,320	1,222	1,222
Year FE	X	X	X	X	X	X	X	X
Employment share, lag	X	X	X	X	X	X	X	X

Source: Authors' calculations.

Note: The table above reports results from regressions of the form:

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \sum_{\tau} \beta_{\tau} \left[\text{Mean Exposure}_{o,T}^{H,\tau} \times \text{Task Share}_{o,T}^{\tau}\right] + c\boldsymbol{\Gamma}_{o,t} + \varepsilon_{o,T}, \tau \in \{M, C, I\}.$$

Panels A, B, and C estimate coefficients using T spanning 1910–2020, 1910–1970, and 1960–2020 (all ranges inclusive), respectively. The variables of interest are Mean Exposure$_{o,T}^{H,\tau}$ $(\tau \in \{M, C, I\})$, technology mean exposure for different task types. Controls $\boldsymbol{\Gamma}_{o,t}$ include year fixed effects and lagged employment share $N_{o,T}$. Columns 1–4 show OLS regressions, while the others show IV regressions. Columns 1, 2, 5, and 6 show regressions with $H = 10$, while the others show regressions with $H = 20$. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

the exposure measure in this fashion to a large extent already captures the concentration of exposure to specific tasks, we report results with and without our baseline concentration measure.

Examining table 4, we see two notable patterns. First, the exposure of manual tasks to technology is negatively predictive of subsequent employment growth at the occupation level consistently. Second, focusing on the entire post-1910 period in panel A, we see that the exposure of cognitive tasks to technology is not significantly related to subsequent employment growth. Panels B and C indicate why this is likely the case: In the 1910–1970 period, the exposure of cognitive tasks to technology is in fact *positively* related to subsequent employment growth. By contrast, during the second period (1960–2020), we see that exposure to cognitive tasks is *negatively* linked with employment growth significantly. Throughout our sample, we see that the exposure of interpersonal tasks to technology—a small share of our overall exposure measure—is not related to subsequent employment growth. Perhaps not surprisingly, decomposing our mean measure into distinct task groups also implies that the relation between our overall task concentration measure and subsequent employment growth is less precisely estimated than before.

Through the lens of the model, these two patterns indicate that technology has consistently substituted for manual tasks throughout the twentieth century. By contrast, in the pre-ICT period, technology served to complement cognitive tasks—in the context of the model, an elasticity of substitution between technology and cognitive tasks that is smaller than the elasticity of substitution across tasks $\nu < \psi$. By contrast, the arrival of computers starting from the 1960s and onward implies that technology improvements such as computer software could also substitute for certain cognitive tasks as well—probably routine cognitive tasks, consistent with the evidence in Kogan and others (2023) and Hampole and others (2025).

HETEROGENEITY BY AGE Although the decennial census provides repeated cross sections, we can track cohorts over time and decompose employment growth by age group. Table 5 reports two sets of longitudinal comparisons. First, we study young workers—age 16–25—over horizons $H \in$ 10, 20:

$$\text{(29)} \qquad \log\left(\frac{N_{o,age\in[16,25],T+H}}{N_{o,age\in[16,25],T}}\right) = \beta_{\text{entrants}} \text{ Mean Exposure}_{o,T}^{H} + \gamma_{\text{entrants}} \text{ Exposure Concentration}_{o,T}^{H} + c_{\text{entrants}} \boldsymbol{\Gamma}_{o,T} + \varepsilon_{o,\text{entrants},T}.$$

Table 5. Technology Exposure and Employment Growth, Age and Cohort Effects

	10-year horizon				*20-year horizon*			
Employment growth, by age and cohort	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
Entrants × mean task exposure	−9.05***	−15.3***	−9.95***	−14.7***	−18.4***	−32.4***	−20.8***	−32.6***
	(1.37)	(2.58)	(1.39)	(2.73)	(2.39)	(4.10)	(2.40)	(4.28)
Entrants × concentration in task exposure		7.40**		5.77*		16.8***		14.6***
		(2.67)		(2.88)		(3.92)		(4.28)
16–25 × mean task exposure	−2.35	−14.6***	−5.18*	−16.1***	−8.76**	−26.6***	−13.4***	−29.3***
	(2.23)	(3.98)	(2.26)	(4.20)	(3.09)	(5.31)	(3.09)	(5.50)
16–25 × concentration in task exposure		14.6***		13.2**		21.4***		19.6***
		(4.05)		(4.35)		(5.43)		(5.62)
26–45 × mean task exposure	−6.70***	−12.4***	−7.28***	−11.9***	−15.4***	−26.2***	−17.0***	−25.7***
	(1.22)	(2.35)	(1.23)	(2.42)	(2.11)	(3.65)	(2.10)	(3.76)
26–45 × concentration in task exposure		6.79**		5.56*		12.9***		10.8**
		(2.25)		(2.39)		(3.66)		(3.84)
46–65 × mean task exposure	−11.4***	−17.3***	−12.1***	−16.9***	−34.7***	−52.5***	−36.1***	−50.3***
	(1.13)	(2.05)	(1.16)	(2.12)	(2.49)	(4.08)	(2.48)	(4.17)
46–65 × concentration in task exposure		6.99***		5.85**		21.2***		17.4***
		(1.94)		(2.09)		(3.94)		(4.13)
Estimator	OLS	OLS	IV	IV	OLS	OLS	IV	IV
N	9,680	9,680	9,680	9,680	9,407	9,407	9,407	9,407
Cohort-year FE	X	X	X	X	X	X	X	X
Cohort × employment share, lag	X	X	X	X	X	X	X	X
Cohort × average occupation age, lag	X	X	X	X	X	X	X	X
p (mid = young, mean)	0.009	0.403	0.213	0.127	0.001	0.898	0.075	0.292
p (mid = young, concentration)		0.002		0.005		0.006		0.004
p (old = young, mean)	0.000	0.415	0.000	0.812	0.000	0.000	0.000	0.000
p (old = young, concentration)		0.019		0.031		0.972		0.639

Source: Authors' calculations.

Note: The table above reports results from regressions of the form:

$$\log\left(\frac{N_{o,age\in[a+H,b+H],T+H}}{N_{o,age\in[a,b],T}}\right)=\beta_{a-b}\text{Mean Exposure}^{H}_{o,T}+\gamma_{a-b}\text{Exposure Concentration}^{H}_{o,T}+c_{a-b}\mathbf{\Gamma}_{o,T}+\varepsilon_{o,a-b,T}$$

for $[a, b] \in \{[16, 25], [26, 45], [46, 65]\}$ (rows 3–8), with the exception of rows 1–2 (entrants), where the dependent variable is $\log\left(\frac{N_{o,age\in[16,25],T+H}}{N_{o,age\in[16,25],T}}\right)$. Controls $\mathbf{\Gamma}_{o,T}$ include year fixed effects, lagged employment share, and lagged average workers age within occupations. The coefficients are estimated separately for each cohort. Coefficients correspond to a one standard deviation change in the independent variable and are multiplied by one hundred. Standard errors (in parentheses) are clustered by occupation. The final four rows of the table display *p*-values on various tests for coefficient equality. ***, **, and * denote statistical significance at the 1 percent, 5 percent, and 10 percent levels, respectively.

Because these workers were at most fifteen in year T, most are new entrants into the labor force. This specification therefore links our exposure measures to entry into exposed occupations. The first rows of table 5 report β_{entrants} and γ_{entrants} across specifications and horizons.

Second, we follow cohorts as they age:

$$(30)\quad \log\left(\frac{N_{o,age\in[a+H,b+H],T+H}}{N_{o,age\in[a,b],T}}\right) = \beta_{a-b}\,\text{Mean Exposure}^{H}_{o,T} + \gamma_{a-b}\,\text{Exposure Concentration}^{H}_{o,T} + c_{a-b}\boldsymbol{\Gamma}_{o,T} + \varepsilon_{o,a-b,T}$$
$$\text{for } [a,b] \in \{[16,25],[26,45],[46,65]\}.$$

This specification measures the effect of exposure on employment growth for young, middle-aged, and older incumbents. Controls include year fixed effects, lagged employment shares, and the lagged average age of workers in the occupation.

Technological transitions may be particularly benign in a world in which reallocation can take place mostly through changes in entry of new workers, as these workers may be most easily able to switch between potential jobs without displacing existing skills and expertise. Stated differently, a model with switching costs that increase with experience (with age acting as a noisy proxy for experience) would likely predict employment elasticities that are declining in age (see, e.g., Kambourov and Manovskii 2009; Cortes and others 2020, for related evidence). If this was indeed the case, we would expect our estimate β_{entrants} to be larger in magnitude than our estimates for overall occupational employment and the analogous coefficient to be smaller for incumbents and decline with age.

Examining table 5, we see that the impact of our exposure measures for entrants is somewhat larger than the pooled coefficients in table 1, which is consistent with the presence of switching costs. Among incumbents, mean exposure coefficients decline sharply with age, with older cohorts significantly more negatively affected. The impact of exposure concentration declines with age, suggesting that younger workers benefit more from opportunities for task reallocation.

Overall, these results show that technology-induced reallocation is not borne mainly by new entrants, but falls heavily on incumbents—especially older workers. The fact that the impact of mean exposure is increasing with age is inconsistent with the idea that switching costs increase with experience—unless these are switches into nonemployment. By contrast, this pattern is consistent with the existence of vintage-specific human capital: New technologies displace existing skills and expertise, and older workers are more likely to hold those vulnerable skills. This evidence complements Kogan and others (2023), who find similar displacement effects in the post-1980 period using data on individual worker earnings.

IV. Implications for the Impact of AI on Labor Demand

Our results so far suggest that the model in section I provides a useful framework for understanding how technological advances shape labor demand. We now turn to the implications of this framework for advances in AI. Forecasting is inherently difficult, so this exercise requires assumptions about which tasks AI is most likely to substitute for, the key model parameters, and the magnitude of spillovers. In section IV.A, we lay out these choices, drawing where possible on the estimates from the previous section. Section IV.B reports the resulting implications for labor demand, and section IV.C discusses the main caveats.

IV.A. Calibration

Here, we briefly discuss our parameterization of the model. We relegate all details to online appendix D.

AI AND WORKER TASKS The first step is to take a stand on which tasks AI is likely to substitute for. In our baseline, we assume AI can perform all cognitive tasks that do not require substantial experience, as measured by the O*NET specific vocational preparation (SVP) score. Tasks requiring less than one year of SVP are treated as fully exposed; tasks requiring one to five years are treated as 50 percent exposed, facing half the capital price shock of a fully exposed task. Tasks are classified into SVP categories using an LLM prompt following Kogan and others (2023), described in the online appendix. Roughly 5 percent of tasks fall in the highest category (more than five years of training) and 32 percent require one to five years, implying that about one-third of tasks are exposed to AI substitution in this baseline. As a robustness check, we also use the task-level generative AI exposure measures from Eloundou and others (2024). We take their preferred β version of exposure, assigning 50 percent exposure to tasks with a score of 0.5 and full exposure to tasks with a score of one.

AI AND PRODUCTIVITY IMPROVEMENTS Next, we take a stand on the decline in the quality-adjusted price of "AI capital." We assume its cost will fall at the same rate as computer hardware. Caunedo, Jaume, and Keller (2023) estimate that computer prices declined by about 14 percent per year between 1984 and 2015. We assume the price of AI capital will decline by the same amount each year over the next decade.

AI AND NEW PRODUCTS AI may also shift labor demand by enabling firms to introduce new products (Babina and others 2024). To capture this channel, we allow for shocks to industry-level shifts in α_I, which we assume are driven by the industry's rate of new AI-related patenting relative to its stock of existing technologies. To simulate industry-specific growth in α_I, we combine the probabilistic assignment of patents to the North American Industry Classification System (NAICS) industries from Goldschlag, Lybbert, and Zolas (2019), data on AI patent counts and growth rates from the United States Patent and Trademark Office AI Patent Database (Pairolero and others 2025), and an estimate of the elasticity of new product creation with respect to patenting from Argente and others (2025).

MODEL PARAMETERS For the key elasticities, we set the elasticity of substitution across occupations, χ, to 1.34 following Caunedo, Jaume, and Keller (2023). In the baseline we also set the elasticity across industries, θ, equal to 1.34 so that spillovers operate only through changes in α_I, and as a robustness check we consider $\theta = 1.72$ from Kogan and others (2023), which introduces spillovers through the industry cost efficiency index Z_I. The elasticity of substitution across tasks is set to $\psi = 0.5$, consistent with Humlum (2021) and Acemoglu and Restrepo (2022), while the degree of decreasing returns at the task level, β, has no clear empirical counterpart; we take $\beta = 0.5$ as the baseline but note that $\beta = 0.75$ yields similar results. The task-level elasticity of substitution between AI capital and labor, ν, is a central parameter in our calibration. We identify it from the empirical relation between employment growth and technology exposure, using the ratio of the estimated β coefficient to the square root of the estimated γ coefficient. The magnitude of this ratio is increasing in ν and independent of the magnitude of the technology shock ε. Targeting the OLS estimates from the ten-year horizon in column 4 of table 2 yields a value of $\nu = 4.63$. Last, we calibrate the across-market elasticity of labor supply ζ to 0.42, following Berger, Herkenhoff, and Mongey (2022).

OTHER ASSUMPTIONS The last step in parameterizing the model requires us to map the observed distribution of employment shares to the model. We do so using the Bureau of Labor Statistics (BLS) Occupational Employment and Wage Statistics (OEWS) data on occupation-industry employment shares, assuming an initial steady state with equal capital intensity across occupations so that wage bill shares match the 2024 distribution. To discipline the remaining task-level CES parameters, we assume symmetric initial capital prices and set the labor share to $s_l = 60$ percent, following Karabarbounis and Neiman (2013).

IV.B. Model Predictions

We focus on the model's predictions for shifts in relative labor demand across occupations. Figure 5 and the first column of table 6 present the baseline results, grouping occupations by educational attainment (panel A), hourly wages (panel B), and female employment share (panel C). Overall, the model implies that AI will raise relative demand for lower-educated, lower-paid, and more male-dominated occupations, partly reversing the patterns of technological change during the twentieth century shown in figure 3.

The effects are economically meaningful. Over the next decade, labor demand for high-education occupations is projected to fall by 0.65 percent per year relative to low-education jobs, though the occupations requiring average levels of education are projected to contract the most. Relative to mid-wage occupations, demand is expected to decline by 0.59 percent annually for managers, 0.29 percent for sales and professionals, and 0.85 percent for clerical and technical jobs. Finally, occupations with higher female shares are predicted to contract relative to male-dominated ones by about 0.53 percent per year.

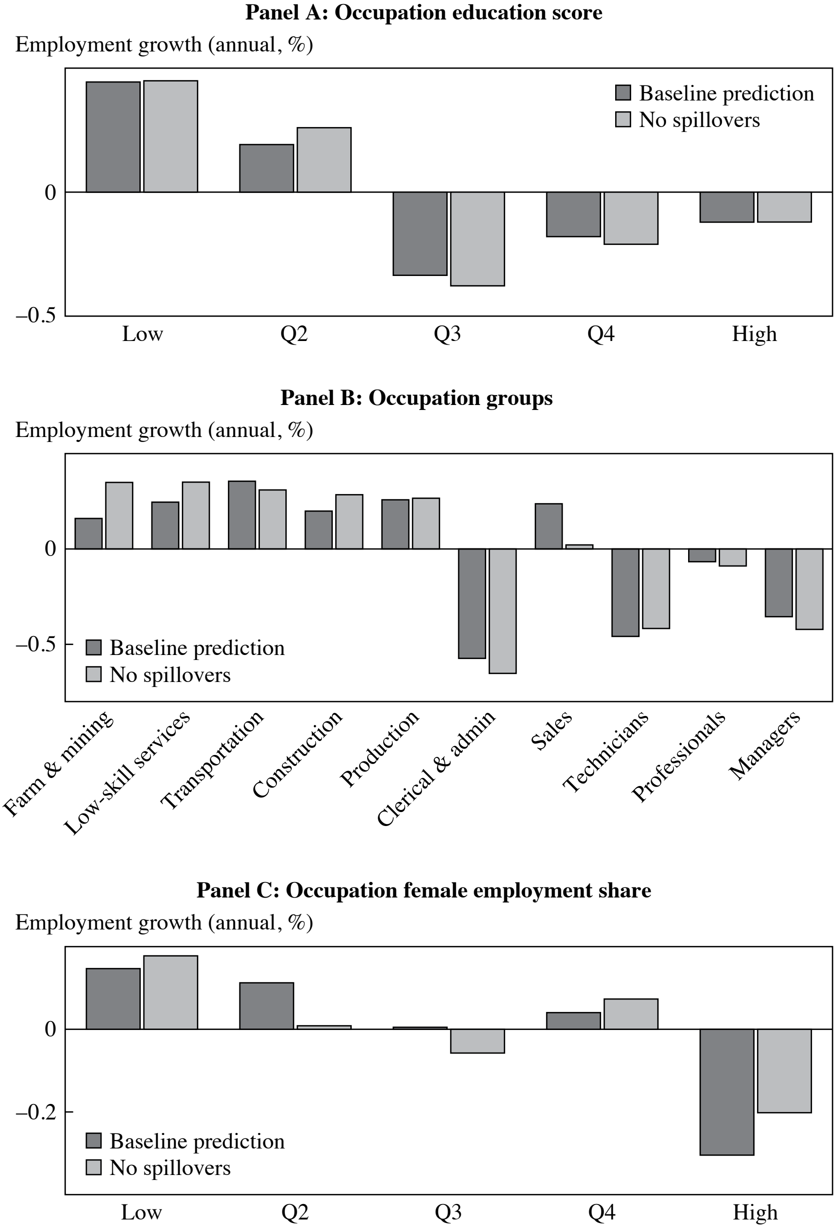
Table 6 demonstrates that these predictions are robust to alternative parameterizations. Eliminating the possibility that AI stimulates new product creation—and hence suppressing industry-level increases in labor demand—yields nearly identical results, as shown in column 2 and figure 5. Column 3 shows that using the task exposure measure from Eloundou and others (2024) produces qualitatively similar shifts, though the magnitudes are notably larger, particularly for sales and professional occupations, which are predicted to face an annual decline in labor demand of 0.82 percent relative to middle-wage jobs. Finally, column 4 indicates that relaxing the assumption $\theta = \chi$ has essentially no effect on the implied reallocation patterns.

IV.C. Caveats

Our simple model makes a number of simplifying assumptions and omits several important economic forces. As such, our analysis comes with several important caveats that we discuss here.

NEW TASKS Advances in AI can lead to the creation of new tasks. This possibility is absent in our model, partly because LLM-sourced task descriptions are too crude to analyze trends in the creation of new tasks.

Figure 5. AI and Shifts in Labor Demand



Source: Authors' calculations.

Note: This figure plots predicted annualized relative employment growth from our AI model simulations described in section IV for occupations categorized by education score, broad occupation group, and share of workers who are female.

Table 6. Model-Implied Impact of AI on Labor Composition

	Model predictions (annualized % rates)			
	Baseline *(1)*	*No product innovation* *(2)*	*Eloundou and others (2024) exposures* *(3)*	*Industry productivity spillovers* $(\theta = 1.72 > \chi = 1.34)$ *(4)*
Panel A: Occupation education level				
High – low	−0.65	−0.71	−1.21	−0.63
Panel B: Occupation wages				
Managers – middle-skill	−0.59	−0.70	−1.05	−0.59
Sales/professionals – middle-skill	−0.29	−0.42	−0.82	−0.28
Clerks/technicians – middle-skill	−0.85	−0.95	−1.12	−0.85
Panel C: Occupation female employment share				
High – low	−0.53	−0.57	−0.56	−0.54

Source: Authors' calculations.

Note: This table examines aggregate outcomes implied by our model simulation of the impact of AI. In the first column we impose our baseline assumption that $\theta = \chi$ and allow for AI-related product innovation. In the succeeding columns we differ from the baseline by dropping product innovation, replacing our baseline assumption that cognitive tasks are exposed to AI with the Eloundou and others (2024) measure of exposure to generative AI and allowing for additional technology spillovers by calibrating $\theta > \chi$ using an estimate for θ taken from Kogan and others (2023). Growth rates are in annualized terms and in percent rates. We assume AI capital prices follow the same average decline as those observed for user cost of computer hardware (13.96 percent per year) over the 1984–2015 period, as calculated by Caunedo, Jaume, and Keller (2023). See section IV and the online appendix for further details.

Depending on which occupations perform these new tasks, this possibility can be an additional driver of cross-occupational shifts in labor demand.

AI AS COMPLEMENT TO LABOR Our model has a sharp distinction between AI capital and human workers. Our baseline assumption is that AI will substitute for certain cognitive tasks that are currently performed by workers. An alternative, perhaps highly speculative, view is that some of these technological advances can enhance the effective quantity of labor. For instance, advances in AI can be similar to improvements in human education—workers could pay a fee to download new skills and expertise immediately to their technology-enhanced brains. In this case, AI can be a labor-augmenting technology that could complement certain types of tasks, similar to the findings for the nonroutine exposure for the 1980–2010 period in Kogan and others (2023).

CAPITAL VERSUS WAGE INCOME The prediction that AI advances are likely to reduce income inequality pertains solely to labor not capital income.

To the extent that financial claims to the new technological advances in AI are not widely owned across households, as in the model of Kogan, Papanikolaou, and Stoffman (2020), advances in AI can lead to an increase in income inequality at the right tail.

SKILL DISPLACEMENT Most importantly, our unit of analysis is a specific occupation, not an individual worker. Even if advances in AI increase demand for a specific occupation, individual workers can be left behind if they lack the skills to fully take advantage of these new technologies—consistent with the evidence in Kogan and others (2023). To the extent that younger workers face a lower cost of acquiring new skills than older workers—or simply have a larger incentive to acquire new skills—we would expect younger workers to be less adversely affected relative to older workers from AI-induced declines in labor demand, consistent with results in table 5.

V. Conclusion

Using new measures of technology exposure that are implied by our model, we document that advances in technology over the last century or more have led to meaningful shifts in labor demand across affected occupations: Labor-saving technology advances have consistently increased the relative labor demand for occupations with higher education requirements, higher wages, and higher fraction of female workers.

ACKNOWLEDGMENTS We are grateful to Matthew Akuzawa and Weizhe Sun for excellent research assistance. We are grateful to William Kerr, Christina Patterson, and Jón Steinsson for helpful comments and suggestions. Dimitris Papanikolaou and Bryan Seegmiller thank the Financial Institutions and Markets Research Center and the Asset Management Practicum for their generous financial support.

References

Abramitzky, Ran, and Leah Boustan. 2017. “Immigration in American Economic History.” *Journal of Economic Literature* 55, no. 4: 1311–45.

Acemoglu, Daron, Ufuk Akcigit, and William R. Kerr. 2016. “Innovation Network.” *Proceedings of the National Academy of Sciences* 113, no. 41: 11483–88.

Acemoglu, Daron, and David H. Autor. 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings.” In *Handbook of Labor Economics*, vol. 4B: 1043–171. Amsterdam: North-Holland.

Acemoglu, Daron, and Pascual Restrepo. 2022. “Tasks, Automation, and the Rise in U.S. Wage Inequality.” *Econometrica* 90, no. 5: 1973–2016.

Aghion, Philippe, Céline Antonin, Simon Bunel, and Xavier Jaravel. 2020. “What Are the Labor and Product Market Effects of Automation? New Evidence from France.” Research Brief 225. Washington: Cato Institute. https://www.cato.org/sites/cato.org/files/2020-07/RB225_final.pdf.

Albanesi, Stefania, and Claudia Olivetti. 2016. “Gender Roles and Medical Progress.” *Journal of Political Economy* 124, no. 3: 650–95.

Argente, David, Salomé Baslandze, Douglas Hanley, and Sara Moreira. 2025. “Patents to Products: Product Innovation and Firm Dynamics.” Working Paper 34592. Cambridge, Mass.: National Bureau of Economic Research.

Aristotle. 1885. *The Politics of Aristotle*, vol. 1, translated by Benjamin Jowett. Oxford: Clarendon Press.

Autor, David H., Caroline Chin, Anna Salomons, and Bryan Seegmiller. 2024. “New Frontiers: The Origins and Content of New Work, 1940–2018.” *Quarterly Journal of Economics* 139, no. 3: 1399–465.

Autor, David H., and David Dorn. 2013. “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market.” *American Economic Review* 103, no. 5: 1553–97.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. “The Polarization of the U.S. Labor Market.” *American Economic Review* 96, no. 2: 189–94.

Autor, David H., and Neil Thompson. 2025. “Expertise.” *Journal of the European Economic Association* 23, no. 4: 1203–71.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson. 2024. “Artificial Intelligence, Firm Growth, and Product Innovation.” *Journal of Financial Economics* 151: 103745.

Bárány, Zsófia L., and Christian Siegel. 2018. “Job Polarization and Structural Change.” *American Economic Journal: Macroeconomics* 10, no. 1: 57–89.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. “Labor Market Power.” *American Economic Review* 112, no. 4: 1147–93.

Berman, Eli, John Bound, and Stephen Machin. 1998. “Implications of Skill-Biased Technological Change: International Evidence.” *Quarterly Journal of Economics* 113, no. 4: 1245–79.

Bose, Gautam, Tarun Jain, and Sarah Walker. 2022. "Women's Labor Force Participation and Household Technology Adoption." *European Economic Review* 147: 104181.

Card, David, and Thomas Lemieux. 2001. "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis." *Quarterly Journal of Economics* 116, no. 2: 705–46.

Caunedo, Julieta, David Jaume, and Elisa Keller. 2023. "Occupational Exposure to Capital-Embodied Technical Change." *American Economic Review* 113, no. 6: 1642–85.

Cortes, Guido Matias, Nir Jaimovich, Christopher J. Nekarda, and Henry E. Siu. 2020. "The Dynamics of Disappearing Routine Jobs: A Flows Approach." *Labour Economics* 65: 101823.

Dauth, Wolfgang, Sebastian Findeisen, Jens Suedekum, and Nicole Woessner. 2021. "The Adjustment of Labor Markets to Robots." *Journal of the European Economic Association* 19, no. 6: 3104–53.

Deming, David J. 2017. "The Growing Importance of Social Skills in the Labor Market." *Quarterly Journal of Economics* 132, no. 4: 1593–640.

de Souza, Gustavo, and Haishi Li. 2023. "Robots, Tools, and Jobs: Evidence from Brazilian Labor Markets." Working Paper 2023-42. Chicago: Federal Reserve Bank of Chicago.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A 'Market Access' Approach." *Quarterly Journal of Economics* 131, no. 2: 799–858.

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock. 2024. "GPTs Are GPTs: Labor Market Impact Potential of LLMs." *Science* 384, no. 6702: 1306–8.

Field, Alexander J. 2003. "The Most Technologically Progressive Decade of the Century." *American Economic Review* 93, no. 4: 1399–413.

Goldin, Claudia. 2006. "The Quiet Revolution That Transformed Women's Employment, Education, and Family." *American Economic Review* 96, no. 2: 1–21.

Goldin, Claudia, and Lawrence F. Katz. 1998. "The Origins of Technology-Skill Complementarity." *Quarterly Journal of Economics* 113, no. 3: 693–732.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race Between Education and Technology*. Cambridge, Mass.: Harvard University Press.

Goldschlag, Nathan, Travis J. Lybbert, and Nikolas J. Zolas. 2019. "Tracking the Technological Composition of Industries with Algorithmic Patent Concordances." *Economics of Innovation and New Technology* 29, no. 6: 582–602.

Goos, Maarten, Alan Manning, and Anna Salomons. 2014. "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring." *American Economic Review* 104, no. 8: 2509–26.

Greenwood, Jeremy, Ananth Seshadri, and Mehmet Yorukoglu. 2005. "Engines of Liberation." *Review of Economic Studies* 72, no. 1: 109–33.

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

Hémous, David, Morten Olsen, Carlo Zanella, and Antoine Dechezleprêtre. 2025. "Induced Automation Innovation: Evidence from Firm-Level Patent Data." *Journal of Political Economy* 133, no. 6: 1975–2028.

Hicks, John. 1932. *The Theory of Wages*. London: Macmillan.

Humlum, Anders. 2021. "Robot Adoption and Labor Market Dynamics." Working Paper, November 28. https://static1.squarespace.com/static/5d35e72fcff15f0001b48fc2/t/61a4e7d3dfd0aa67849b9918/1638197204858/humlumJMP.pdf.

Jerome, Harry. 1934. *Mechanization in Industry*. Cambridge, Mass.: National Bureau of Economic Research.

Jiang, Wei, Junyoung Park, Rachel (Jiqiu) Xiao, and Shen Zhang. 2025. "AI and the Extended Workday: Productivity, Contracting Efficiency, and Distribution of Rents." Working Paper 33536. Cambridge, Mass.: National Bureau of Economic Research.

Jiang, Wei, Yuehua Tang, Rachel J. Xiao, and Vincent Yao. 2025. "Surviving the Fintech Disruption." *Journal of Financial Economics* 171: 104071.

Kambourov, Gueorgui, and Iourii Manovskii. 2009. "Occupational Specificity of Human Capital." *International Economic Review* 50, no. 1: 63–115.

Karabarbounis, Loukas, and Brent Neiman. 2013. "The Global Decline of the Labor Share." *Quarterly Journal of Economics* 129, no. 1: 61–103.

Katz, Lawrence F., and Robert A. Margo. 2014. "Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective." In *Human Capital in History: The American Record*, edited by Leah Platt Boustan, Carola Frydman, and Robert A. Margo: 15–57. Chicago: University of Chicago Press.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107, no. 1: 35–78.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. 2021. "Measuring Technological Innovation over the Long Run." *American Economic Review: Insights* 3, no. 3: 303–20.

Keynes, John Maynard. 1978. "Economic Possibilities for Our Grandchildren (1930)." In *The Collected Writings of John Maynard Keynes, Volume 9: Essays in Persuasion*, edited by Elizabeth Johnson and Donald Moggridge: 321–32. London: Royal Economic Society.

Koch, Michael, Ilya Manuylov, and Marcel Smolka. 2021. "Robots and Firms." *Economic Journal* 131, no. 638: 2553–84.

Kogan, Leonid, Dimitris Papanikolaou, Lawrence D. W. Schmidt, and Bryan Seegmiller. 2023. "Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data." Working Paper 31846. Cambridge, Mass.: National Bureau of Economic Research.

Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman. 2020. "Left Behind: Creative Destruction, Inequality, and the Stock Market." *Journal of Political Economy* 128, no. 3: 855–906.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68, no. 5: 1029–53.

Mann, Katja, and Lukas Püttmann. 2023. "Benign Effects of Automation: New Evidence from Patent Texts." *Review of Economics and Statistics* 105, no. 3: 562–79.

Manyika, James, Susan Lund, Michael Chui, Jacques Bughin, Lola Woetzel, Parul Batra, Ryan Ko, and Saurabh Sanghvi. 2017. *Jobs Lost, Jobs Gained: What the Future of Work Will Mean for Jobs, Skills, and Wages*. Washington: McKinsey Global Institute.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. "Distributed Representations of Words and Phrases and Their Compositionality." Working Paper, October 16. arXiv:1310.4546.

Ngai, L. Rachel, Claudia Olivetti, and Barbara Petrongolo. 2024. "Gendered Change: 150 Years of Transformation in US Hours." Working Paper 32475. Cambridge, Mass.: National Bureau of Economic Research.

Ngai, L. Rachel, and Barbara Petrongolo. 2017. "Gender Gaps and the Rise of the Service Economy." *American Economic Journal: Macroeconomics* 9, no. 4: 1–44.

O'Rourke, Kevin H., Alan M. Taylor, and Jeffrey G. Williamson. 1996. "Factor Price Convergence in the Late Nineteenth Century." *International Economic Review* 37, no. 3: 499–530.

Pairolero, Nicholas A., Alexander V. Giczy, Gerard Torres, Tisa Islam Erana, Mark A. Finlayson, and Andrew A. Toole. 2025. "The Artificial Intelligence Patent Dataset (AIPD) 2023 Update." *Journal of Technology Transfer* 50, no. 6: 2587–610.

Parsons, Donald O. 1980. "The Decline in Male Labor Force Participation." *Journal of Political Economy* 88, no. 1: 117–34.

Webb, Michael. 2020. "The Impact of Artificial Intelligence on the Labor Market." Working Paper, January. https://www.michaelwebb.co/webb_ai.pdf.

Comments and Discussion

COMMENT BY

WILLIAM KERR Liu, Papanikolaou, Schmidt, and Seegmiller tackle in this paper a first-order question in macro labor: how technological innovations, filtered through the lens of tasks to be performed in the workplace, shape occupational employment over the long run, and what the historical record implies for the potential labor impacts of artificial intelligence (AI). The authors' contribution is threefold. First, they offer a task-based model that delivers sufficient statistics—a "mean exposure" term that captures the total intensity with which a technology overlaps an occupation's task set, and an "exposure concentration" term that captures how unevenly that exposure is distributed across tasks—so that empirical work maps directly to theory. Second, they construct a frontier long-run measurement framework that uses large language model (LLM) text methods to link patents to occupation-specific tasks over more than a century. Third, they provide empirical evidence that higher mean exposure predicts subsequent declines in occupational employment, while (conditional on mean exposure) greater exposure concentration mitigates those declines by enabling within-occupation reallocation toward nonexposed tasks. The authors conclude with a forward-looking calibration of the labor impacts of AI, which suggests a future retilting of relative labor demand across education and occupation groups. Whereas recent technologies have typically favored highly educated and skilled workers, that may be less likely going forward given AI's capabilities to take on advanced cognitive tasks.

CORE PAPER There is much to admire in the core paper, especially its tight integration of theory, measurement, and evidence. The model provides

Brookings Papers on Economic Activity, Fall 2025: 273–292

sufficient statistics that the empirical work estimates directly, confirming a negative impact on mean exposure and a positive coefficient on exposure concentration. Built around decadal data, the long-run perspective also adds value by showcasing technology eras. The authors have the capability to separate patent volume and task exposure, showing for example that many chemical and electronics breakthroughs generated abundant patents with little direct task exposure, while woodworking, railroads, and later software had a disproportionate impact on some tasks and thus occupations. This paper thus brings new micro data sources to issues often discussed economy-wide in theory and measured in aggregates. Heterogeneity analyses by education, wages, and gender indicate a tilt of demand toward higher-education and higher-wage occupations on average and those staffed disproportionately by female workers. Some wedges narrow post-1980 as information and communications technology begins to substitute for more cognitive tasks. The work provides some intriguing evidence that job polarization may have commenced by the 1960s or before, earlier than often thought.

The authors' use of powerful frontier LLMs for data generation and measurement makes all this possible. Different models pursue the construction of likely task descriptions of occupations for earlier eras (since the *Dictionary of Occupational Titles* was first published in 1939) and compare the text of patents during eras to determine potential automation. Each step appears sound, but there may be unanticipated ways that the LLM language patterns cast today's information backward and possibly create circular links. To their credit, the authors are experts in this frontier work and use alternative LLMs, but differences are larger in the nineteenth century, precisely where measurement is hardest. As they take the data forward for further uses, I encourage the authors to run more adversarial models (or use good old-fashioned human grading, lest technology take on all roles!) that explicitly test historical feasibility (e.g., "Does this set of tasks accurately characterize this occupation's work in 1880? Which task looks most anomalous?"). Confidence thresholds brought to the generated data would confirm and clarify their information signal.

There are also immediate avenues for the empirical work to take the next step. The authors tackle the direction of technology change, but they mostly forgo, for now, the pace of change. This rate would be quite material for thinking through, for example, the ability of workers to reskill over tasks, in addition to concomitant productivity enhancements. The framework also abstracts from the required diffusion of newly patented technologies to the workplace, which takes a while and has been at different paces over the 160 years spanned by this study. This relates to a need to unpack the serial and cross-correlations of the mean task exposure of an occupation and its exposure concentration, which are tightly connected to each other as shown in figure 2 and table 2 of the paper. A dutiful discussion about a technology paper should also (in all times and places) advocate for nonpatent alternatives for technology to be incorporated into an analysis, but that is far easier said from the sidelines than accomplished in practice.

CONCEPTUAL CONSIDERATIONS ABOUT OCCUPATIONS The heart of my reflections on this paper, however, revolve around the strengths and liabilities of using an occupation-task framework to investigate labor impacts from technology change. The theory behind this research stream makes the captured channels clear and accessible, and they also can be used to ponder what is missing. The tight link of the theory and empirics, which I praised above and connects to Hampole and others (2025), means that we need to contemplate the framework seriously to understand where additional impact for workers from technologies could come about.[1]

The lens of tasks is unquestionably discerning in many settings. Chipotle founder Steve Ells opened Kernel in New York City in February 2024 with the ambition of running the fast casual restaurant with three people compared to a typical team of ten or more (Kingson 2024). The description of automation undertaken at Kernel was almost one for one with tasks: A six-axis robot arm portioned and fed food items onto a conveyor for reheating, ordering and pickup were fully conducted through digital apps and lockers, and so on. The venture turned out not to be successful, but the authors' model captures the ambition exactly, including the consolidation of taskwork for those who remained at the restaurant to nonautomated roles. Many images of technology in the workplace—from more automated warehouses to virtual bots for administrative work—are similarly intuitive in this approach.

A second step requires thinking through the structure and boundaries embedded in the occupational classifications collected with the decennial

1. Conceptually, the task-based approach echoes and extends the framework in Acemoglu and Autor (2011) and the broader task-centric narrative that grows out of Autor, Levy, and Murnane's (2003) seminal work. Prominent empirical research strands include routine-biased technological change and polarization (Autor and Dorn 2013; Michaels, Natraj, and Van Reenen 2014; Jaimovich and Siu 2020), robots and local labor market effects (Graetz and Michaels 2018; Acemoglu and Restrepo 2020; Dauth and others 2021), within-occupation task change captured in vacancies and historical job ads (Hershbein and Kahn 2018; Atalay and others 2020), early AI exposure indexes (Felten, Raj, and Seamans 2021; Webb 2020), and field evidence on LLMs (Brynjolfsson, Li, and Raymond 2025; Noy and Zhang 2023).

census. Looking at the underlying codes provides many assurances. For example, technology observers often comment on the cataclysmic failure of video rental stores in the face of streaming. But the occupational code in question is "retail sales clerk," which captures well the capacity of potentially displaced individuals from the shuttered Blockbuster to easily find new work at other retail establishments. But there are also trickier examples. One of the occupations is "miner," which carries lots of political significance, to put it mildly. The classification is not "energy extraction worker." Thus, the task-occupation framework would capture well efforts by technology to enhance the productivity of a coal miner, while missing that the most significant innovations for the fate of a coal miner sit in other sectors, such as the rapid improvements in solar power technologies and battery storage, which can substitute for coal as an energy input.

Stretching slightly the concept of "concentration" that is central to the authors' work, we can also recognize additional ways technology, tasks, and occupations interact, which are less captured by the model. In the medical profession, scans of patients via CT (computed tomography) and MRI (magnetic resonance imaging) machines are important and in high demand. These machines are very expensive capital investments, and hospitals and practices seek to continually utilize them. Medical professionals help patients into and out of the machines, and others run and analyze the scans. IONIC Health, a Brazilian startup recently acquired by GE HealthCare (2023), saw an opportunity in the amount of waiting time that individuals running the machines experienced. The company developed new technology that retrofits traditional machines to allow remote operation: remote operational control of the machine itself, necessary data feeds, and video links. By itself, this technology could result in labor cost savings by placing the individuals running the equipment in less expensive locations. The bigger impact for labor, though, came from the new capacity of a dedicated external employee to manage multiple machines at different locations in parallel, overcoming the idle time associated with a single machine and thus significantly boosting labor productivity. Similar consolidations, often combined with remote work, have occurred in port management and manufacturing, and remote surgery may be a future use case.

Agrawal, Gans, and Goldfarb (2018) also contemplate two other types of consolidations worth dwelling upon. The first recognizes the incentives to "close the loop" on technology's role in a workplace or work stream when full automation becomes feasible. Imagine a warehouse where every task is automated except for a final task that remains conducted by humans. There is significant and disproportionate incentive to automate that last step because it thereby frees the overall operation of the warehouse in many respects, such as its physical location and internal layout. In one heavily automated facility I visited, the stored goods no longer needed to be grouped together by type or even by owner, as the system had complete awareness and the collecting robots whizzed to where each item was located. The facility was several stories tall, but most items were stored on deep shelves that were not easily accessible to workers. I earlier commended the authors for having a model and empirical approach that could be linked to micro data to see impacts of technology in the workplace. Future work could build upon the distinctions of mean exposure versus exposure concentration to capture these settings.

It will also be fascinating to observe how occupational definitions change with new technologies and reallocated work. Consider the classic example of a school bus driver. When autonomous vehicles take on the core task of the bus driver's job, most parents are still going to want to have a human on board the school bus, if for no other reason than to monitor the kids and ensure their safety. If not driving the bus, then what will the employee also do? Perhaps they start a class early, converging to a traditional "teacher" job title. Or maybe the work gets fully reoriented in unanticipated ways. The authors' data can shed insight into these occupational adjustments in the face of technological changes, similar to Lin's (2011) work tracing the emergence of new occupations to more diversified cities.

To close, it is useful to note empirical evidence on the importance of declining occupations, which sits at the heart of this work and Kogan and others (2023). This paper estimates relative occupational demand, not the employment and earnings trajectories of workers employed in those occupations. Policymakers also care deeply about how workers experience these adjustments, which includes the capacity of displaced workers to reposition themselves in the labor market and undertake reskilling. Building on a large literature that considers job loss and entry into weak labor markets (Jacobson, LaLonde, and Sullivan 1993; Davis and von Wachter 2011; Kahn 2010; Oreopoulos, von Wachter, and Heisz 2012), Edin and others (2023) quantify the impact of being in a declining occupation in Sweden during the 1986–2013 period; ongoing work by Barth and others (2026a, 2026b) explores parallel patterns in Norwegian data for 2007–2024 and in US data for 2000–2020. In the Swedish experience, workers initially employed in declining occupations (measured as a 25 percent net decline in employment in the occupation over the period) show diminished careers. Conditional on many individual traits including initial earnings, workers in declining occupations suffer about 2.1 percent fewer work years during the ensuing period and experience about 5.0 percent lower career earnings. Larger earnings losses, amounting to about 10 percent, can occur for those starting at the bottom of the occupation-specific wage distribution.

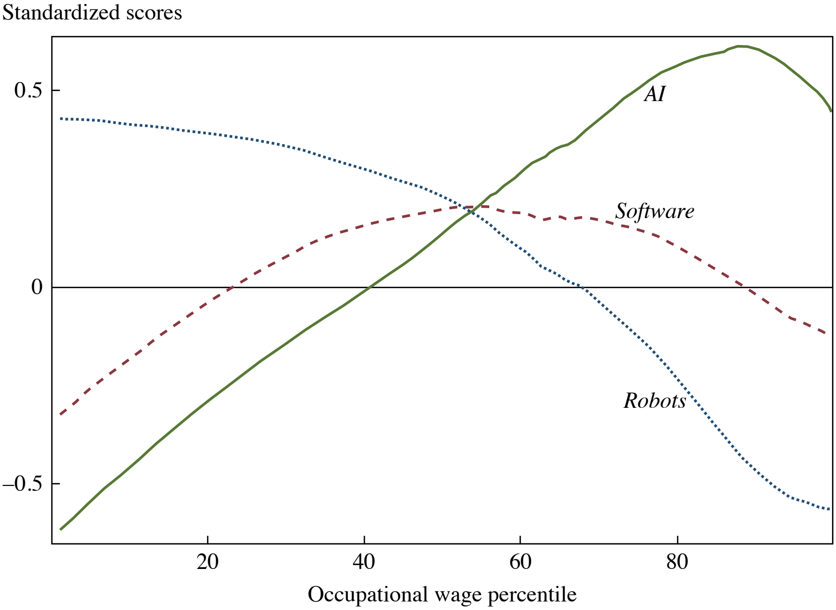
LOOKING AHEAD TO AI'S IMPACT The authors close with provocative forecasts of how AI will have an impact on the labor market based upon tasks where we have seen past and ongoing use of the new technology. This approach leads to a reasonable conclusion that AI's impact will differ from those of recent technology waves that mostly served to heighten demand for and compensation of highly educated and skilled workers relative to the rest of the labor force. If AI increasingly takes on the tasks of these high-skilled workers less exposed to prior waves, labor demand may shift in relative terms over coming years toward occupations that have traditionally been lower-paid and disproportionately staffed by lower-educated male workers.

There is a trickiness, to put it mildly, in modeling the likely consequences of seismic shifts in technology, which AI appears to be initiating. Tech futurists like Lee (2018) worry that a task-oriented model misses the scope for sector-killing innovations, but the worry can be even more mundane. Dial back to the mid-2000s and consider the advent of the smartphone, itself building on the technology wave of the Internet and World Wide Web. Forward-looking pundits might well have picked some tasks the technology could one day be used for, such as customer interfaces, payments, and navigation. But even if aware of the launch of UberCab in March 2009, few of us would have contemplated that ride-sharing would be the largest consequence (thus far) for the workforce from smartphones. The tasks can be visible, but the combinations and integrations of them are nearly impossible to forecast. To overly bake this point, oDesk (later to become Upwork) was founded in 2003 to provide global contracting of digital work via the Internet. I would have bet quite a bit that the potential of placing almost any digital office task anywhere in the world would have swamped the impact of ride-sharing. In fall of 2025, Uber's market cap was $208 billion, compared to $2.1 billion for Upwork.[2]

Returning to academics, Webb (2020) gives one of the first representations of how AI technologies affected different parts of the occupational wage distribution using the text of patents and occupational tasks. His

2. Market capitalizations from November 3, 2025, using https://public.com/stocks/upwk/market-cap and https://public.com/stocks/uber/market-cap.

Figure 1. Impact of Patented Technologies on Tasks by Occupational Wage Percentile



Source: Reproduced from presentation material by Webb (2020) with author's permission.

descriptive work showcases three waves of technologies. Figure 1, prepared by Webb for presentations, shows the waves together using standardized scores. Robotics held significant impact for occupations in the bottom half of the wage distribution, commensurate with its historical consequences for sectors like agriculture and manufacturing. Computers and software held significant impact for occupations sitting in the middle of the wage distribution, closely aligned with the polarization literature that captures recent labor market dynamics. While AI was still emerging at the time of Webb's study, it showed its most significant impacts in the upper part of the wage distribution, peaking near to the 90th percentile. This aligns with many of the early use cases for AI, such as marketing.

The authors' projections in this paper of the impact of AI on the labor market are closely aligned with Webb's representation. My somewhat unfair and orthogonal critique is that significant investment is being made to link AI with robotics and related technologies that bring the skills of AI into the physical world. The robotics arms in restaurants like Kernel will benefit from learning from other robotic arms. A world full of autonomous vehicles may constitute the ultimate expression of AI given that what is needed for self-driving cars is the ability to handle nonroutine cognitive issues facing drivers. If I view AI's impact in isolation and only cognitive in nature, I agree the impacts sit at the upper end of the occupational wage distribution. If we open the aperture to consider its codevelopment with robotics, all bets are off for me as to which occupations will face the greatest substitution threats at the task level versus complementary reinforcement.

REFERENCES FOR THE KERR COMMENT

Acemoglu, Daron, and David H. Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, vol. 4B: 1043–171. Amsterdam: North-Holland.

Acemoglu, Daron, and Pascual Restrepo. 2020. "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy* 128, no. 6: 2188–244.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb. 2018. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston: Harvard Business Review Press.

Atalay, Enghin, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. 2020. "The Evolution of Work in the United States." *American Economic Journal: Applied Economics* 12, no. 2: 1–34.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103, no. 5: 1553–97.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118, no. 4: 1279–333.

Barth, Erling, Maria Forthun Hoen, Sari Pekkala Kerr, and William R. Kerr. 2026a. "Declining Occupations and Career Outcomes in Norway." Working Paper 35096. Cambridge, Mass.: National Bureau of Economic Research. (Forthcoming in *AEA Papers and Proceedings*.)

Barth, Erling, Maria Forthun Hoen, Sari Pekkala Kerr, and William R. Kerr. 2026b. "When Is Mobility Possible? Escaping from Declining Occupations." Working Paper (unpublished).

Brynjolfsson, Erik, Danielle Li, and Lindsey Raymond. 2025. "Generative AI at Work." *Quarterly Journal of Economics* 140, no. 2: 889–942.

Dauth, Wolfgang, Sebastian Findeisen, Jens Suedekum, and Nicole Woessner. 2021. "The Adjustment of Labor Markets to Robots." *Journal of the European Economic Association* 19, no. 6: 3104–53.

Davis, Steven J., and Till von Wachter. 2011. "Recessions and the Costs of Job Loss." *Brookings Papers on Economic Activity*, Fall: 1–55.

Edin, Per-Anders, Tiernan Evans, Georg Graetz, Sofia Hernnäs, and Guy Michaels. 2023. "Individual Consequences of Occupational Decline." *Economic Journal* 133, no. 654: 2178–209.

Felten, Edward, Manav Raj, and Robert Seamans. 2021. "Occupational, Industry, and Geographic Exposure to Artificial Intelligence: A Novel Dataset and Its Potential Uses." *Strategic Management Journal* 42, no. 12: 2195–217.

GE HealthCare. 2023. "GE HealthCare Announces New Version of Digital Expert Access, the First FDA 510(k)-Cleared Device to Enable Remote Patient Scanning, and Exclusive Distribution Agreement with IONIC Health." Press Release, November 21. https://investor.gehealthcare.com/news-releases/news-release-details/ge-healthcare-announces-new-version-digital-expert-access-first.

Graetz, Georg, and Guy Michaels. 2018. "Robots at Work." *Review of Economics and Statistics* 100, no. 5: 753–68.

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

Hershbein, Brad, and Lisa B. Kahn. 2018. "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings." *American Economic Review* 108, no. 7: 1737–72.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review* 83, no. 4: 685–709.

Jaimovich, Nir, and Henry E. Siu. 2020. "Job Polarization and Jobless Recoveries." *Review of Economics and Statistics* 102, no. 1: 129–47.

Kahn, Lisa B. 2010. "The Long-Term Labor Market Consequences of Graduating from College in a Bad Economy." *Labour Economics* 17, no. 2: 303–16.

Kingson, Jennifer A. 2024. "Inside the Chipotle Founder's New Robot-Powered Vegan Restaurant." Axios, March 11. https://www.axios.com/2024/03/11/chipotle-founder-kernel-robots-vegan-new-restaurant.

Kogan, Leonid, Dimitris Papanikolaou, Lawrence D. W. Schmidt, and Bryan Seegmiller. 2023. "Technology and Labor Displacement: Evidence from Linking Patents with Worker-Level Data." Working Paper 31846. Cambridge, Mass.: National Bureau of Economic Research.

Lee, Kai-Fu. 2018. *AI Superpowers: China, Silicon Valley, and the New World Order*. Boston: Houghton Mifflin Harcourt.

Lin, Jeffrey. 2011. "Technological Adaptation, Cities, and New Work." *Review of Economics and Statistics* 93, no. 2: 554–74.

Michaels, Guy, Ashwini Natraj, and John Van Reenen. 2014. "Has ICT Polarized Skill Demand? Evidence from Eleven Countries over 25 Years." *Review of Economics and Statistics* 96, no. 1: 60–77.

Noy, Shakked, and Whitney Zhang. 2023. "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence." *Science* 381, no. 6654: 187–92.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. 2012. "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics* 4, no. 1: 1–29.

Webb, Michael. 2020. "The Impact of Artificial Intelligence on the Labor Market." Working Paper, January. https://www.michaelwebb.co/webb_ai.pdf.

COMMENT BY

CHRISTINA PATTERSON The central question that Liu, Papanikolaou, Schmidt, and Seegmiller study—whether new technologies ultimately displace workers or create new opportunities—is almost as old as economics itself. David Ricardo (2004, p. 388) worried in the early nineteenth century that "the substitution of machinery for human labour" could be "very injurious" to workers, and a century later John Maynard Keynes (1978) warned of looming "technological unemployment." It seems fitting, therefore, that historical data spanning centuries would be a powerful place to look for answers to this question. This paper takes an ambitious step in that direction, using recent advances in large language models (LLMs) to build a comprehensive and unified historical data set of both technologies and the tasks or industries that they affect.

Given that this paper tackles a classic question, it is worth reflecting on how it relates to and extends an already substantial body of work documenting that technology reshapes labor demand. The literature on skill-biased technological change shows that new technologies have historically raised the relative demand for high-skill workers, helping to explain rising wage inequality (Katz and Murphy 1992; Autor, Katz, and Kearney 2006). The capital-skill complementarity literature offers a complementary view: New forms of capital tend to complement skilled labor more than unskilled labor, leading again to a reallocation of labor demand and a higher skill premium (Krusell and others 2000; Goldin and Katz 1998, 2008). More recently, task-based frameworks have shifted the focus from skills to conceptualizing jobs as bundles of tasks: some routine and automatable, others nonroutine and hard to codify. This work has been central to our understanding that technology causes not just a bias toward higher-skill jobs but also job polarization and the hollowing out of middle-skill routine jobs (Autor, Levy, and Murnane 2003; Acemoglu and Autor 2011; Autor and Dorn 2013).

We know much less about the channels through which specific technologies affect labor demand. This paper adds to our understanding of these questions by interpreting an enormous amount of historical data through the lens of a structural model. The model builds on the task-based frameworks, specifically that of Hampole and others (2025), and conceptualizes the effect of technology on labor demand as coming through three terms. First, labor demand for an occupation depends on the average exposure of an occupation's tasks to task-related technologies. In principle, such technologies could be labor-augmenting (e.g., a typewriter for an author) or labor-saving (e.g., a loom for a weaver), but empirically, the authors show that most of the task-related technologies over the past several decades have been labor-saving. Second, labor demand depends on how unevenly exposure to new technologies is distributed across an occupation's tasks. If exposure is concentrated on a few tasks, there may be scope for workers to reallocate effort to other, less affected tasks within the same occupation. Third, labor demand is affected by spillovers, capturing how technological innovations affecting some tasks in an industry feed into productivity growth and product demand more broadly.

This structure allows the authors to bring a long-run perspective to familiar themes in the literature on skill-biased and routine-biased technological change. Broadly speaking, the historical patterns they uncover line up well with what we have come to expect: Technology has contributed to employment polarization, with disproportionate exposure of middle-skill occupations, and this process appears to have begun earlier than the standard post-1980 narrative suggests. They also find that technology has consistently favored higher-education, higher-wage, and more female occupations over much of the twentieth century, in line with work such as Goldin and Katz (2008) and Autor, Katz, and Kearney (2006). In this sense, the main contribution is less to overturn conventional wisdom than to place it on a firmer and more unified empirical footing, showing how a single framework can organize a wide range of historical facts about technology and labor demand.

The rest of my comments focus on the authors' interpretation of the economic forces driving the empirical patterns that they recover. In particular, I discuss what we learn from the "concentration" and "industry spillover" terms, and how we should interpret these historical findings when thinking about the consequences of the new wave of disruptive technologies from artificial intelligence (AI).

INTERPRETING THE CONCENTRATION TERM The concentration term in the model is fundamentally about reallocation potential within an occupation. If a new technology affects only a few tasks in an occupation, workers can shift effort to other tasks that become relatively more valuable and this dampens the drop in labor demand for those workers.

Mathematically, this force appears as a variance of exposure across tasks, weighted by each task's importance in production. Tasks that account for a large share of the occupation's contribution to output receive more weight, while marginal tasks receive relatively little. A high weighted variance means that exposure is unevenly distributed across economically important tasks.

In the empirical implementation, however, the concentration term is computed as the *unweighted* variance of exposure across the tasks that the

LLM produces for each occupation and decade. This is a reasonable and transparent starting point, given that historical data do not provide task-level production weights, but it introduces a subtle and important question: What counts as a task? The paper relies on an LLM to generate task lists, and in practice this consistently yields ten to fifteen tasks per occupation per decade. Take, for example, a telephone operator in the 1920s. When prompted, the LLM provided me a list of twelve tasks: plugging cords, monitoring lights, logging long-distance calls, ringing subscribers, providing directory assistance, routing emergency calls, and so on. From a production perspective, many of these are not distinct tasks. Instead, they seem both to me and ChatGPT (when prompted differently) to be fragments of one core task, which one might summarize as "manually connecting and managing calls."

This seems to be a general pattern, not one specific to the selected case of a telephone operator. For example, occupations such as tollbooth operators and dentists often end up with a similar number of tasks, even though one occupation arguably has many more truly distinct tasks than the other. This suggests that the LLM has limited discipline on task aggregation—core tasks tend to be split into several subtasks to reach a list of ten to fifteen tasks for each occupation-decade.

This is an important distinction when tasks are not weighted in the construction of both the variance and average task exposure terms. Consider what happened with introduction of the automatic telephone exchange system, which mechanically connected calls and dramatically reduced the need for human operators. In the authors' framework, such a technology would likely appear as affecting many of the LLM-generated subtasks: plugging cords, watching lights, ringing phones, disconnecting calls, and monitoring call quality. From the perspective of the unweighted variance of exposure, this looks like a diffuse pattern: Many tasks are hit, so exposure appears spread out, and concentration appears low. But conceptually, this is a technology that annihilated a single core task. By contrast, if a technology mainly touched a narrow, peripheral set of tasks, then exposure could look highly concentrated. Indeed, it seems plausible that the methodology is most likely to yield concentrated exposure when it is primarily peripheral tasks that are affected.

This observation suggests two different narratives behind the strong positive coefficient on the concentration term in the authors' table 1. The first is the one implied by the model: When exposure is concentrated in a subset of tasks, workers can reallocate within the occupation to tasks that are less affected, and this reallocation quantitatively mitigates the employment losses from the technology. An alternative interpretation, however, is that smaller employment losses occurred when core tasks were spared. Under this view, the positive coefficient on concentration is picking up the fact that employment falls much more when a core task is affected than when only peripheral tasks are affected. This would be true even if there was no reallocation of workers across tasks.

These two interpretations have very different implications. The first tells an optimistic story in the face of AI: Even if AI automates some core tasks within an occupation, workers may reallocate within the job, and the adverse employment effects may be muted. The second paints a more fragile picture: Resilience depends on core tasks being left intact. My sense is that reality probably combines elements of both, and disentangling them is important if we want to extrapolate the historical evidence to the future.

One constructive way forward would be to bring information on task importance into the empirical concentration measure. The authors note that they do not have direct data on $\alpha(j)$, the production weights, but the spirit of the paper is to lean on modern language models when traditional data are unavailable. The authors could ask the LLM not only to list tasks but also to assign approximate importance weights—for example, the share of time or value added associated with each task. In my telephone operator experiment, the LLM suggested to me that roughly 80 percent of the job was in the core activity of connecting and monitoring calls. Using such weights to construct a weighted variance of exposure would bring the empirical measure closer to the model's sufficient statistic and help clarify whether the large positive concentration coefficient is really about reallocation or about which tasks are affected.

INTERPRETING INDUSTRY SPILLOVERS My second comment concerns the interpretation of the industry spillover term. In the model, industry spillovers arise from two sources. First, they come from spillovers in the true sense of the word, which occur when task-specific technologies raise industry-level productivity, lower effective production costs, and thereby increase demand for the industry's output. That, in turn, raises labor demand for *all* occupations in the industry, including those not directly automated. Second, this term captures total factor productivity shocks to an industry, which occur potentially independently from the labor-related technologies.

To highlight the relevance of these alternative technologies, let us return to the telecommunications industry in the early twentieth century. For telephone operators, the automatic telephone exchange was a direct, task-specific innovation that automated the core task of connecting calls. Feigenbaum and Gross (2024) show that its adoption led to large and rapid job losses for operators. At the same time, there were other important innovations in the industry, such as long-distance repeaters that extended the reach of calls, cables that increased line capacity, improved circuits that reduced noise, and desk sets that made telephones more attractive for households. All of these expanded the market for telephone service and increased demand for the industry.

All of these innovations enter the empirical industry spillover measure through the same channel, in that they all reflect patenting growth in the telecommunications industry. Thus, when we see that higher industry patenting is associated with higher employment across occupations, we learn that industries undergoing rapid innovation tend to experience stronger labor demand growth. It does not directly inform the strength of the spillovers.

A second comment, closely related, is that because the industry spillover term is specified at the industry level and enters the regression uniformly for all occupations in that industry, it may mask within-industry polarization. In the telecommunications example, some innovations (like long-distance repeaters) likely increased demand for both operators and technical workers. However, that technology likely had no offsetting demand effect for the telephone operators, who were no longer needed for local calls after the mechanical switch. In the empirical specification, however, the positive industry patent term shifts employment for all telecommunications occupations in the same direction, effectively assigning local operators some of the demand growth generated elsewhere in the industry. This uniform spillover therefore tends to compress differences in employment outcomes across occupations within an industry. My sense is that, absent this averaging force, the authors' estimates of the role of technology in driving polarization across occupations would be even stronger.

FROM JOBS TO WORKERS A final observation, which the authors themselves emphasize, is that the analysis is conducted at the level of the occupation (or the occupation by industry). That is a reasonable first step in an analysis that spans nearly two centuries of data. However, it is important to remember that what we care about is ultimately the impact of technology on workers, not just jobs.

The telephone operator case is again a helpful reminder of the importance of this distinction. Automation effectively eliminated the occupation in many firms, but worker-level outcomes were more muted because workers reallocated across occupations over time. Some younger cohorts never entered the occupation in the first place, some older workers transitioned into clerical or service roles, and others left the labor force. Occupation-level employment data, even with excellent measures of technological exposure, cannot speak to the importance of this cross-occupation reallocation. Understanding the impact of these technologies on workers is an important next step.

CONCLUSION This paper uses new methods to construct a sweeping data set of occupation-level exposure to technological innovations over the past several decades. The model provides a uniform framework to interpret these historical data, and the evidence confirms that technology has both direct and offsetting effects on labor demand. Understanding the balance between these forces is essential for thinking about workers' prospects in the upcoming era of rapid innovation, and I look forward to seeing work that builds on these data to continue to learn from the past.

REFERENCES FOR THE PATTERSON COMMENT

Acemoglu, Daron, and David H. Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, vol. 4B: 1043–171. Amsterdam: North-Holland.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103, no. 5: 1553–97.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *American Economic Review* 96, no. 2: 189–94.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *Quarterly Journal of Economics* 118, no. 4: 1279–333.

Feigenbaum, James, and Daniel P. Gross. 2024. "Answering the Call of Automation: How the Labor Market Adjusted to Mechanizing Telephone Operation." *Quarterly Journal of Economics* 139, no. 3: 1879–939.

Goldin, Claudia, and Lawrence F. Katz. 1998. "The Origins of Technology-Skill Complementarity." *Quarterly Journal of Economics* 113, no. 3: 693–732.

Goldin, Claudia, and Lawrence F. Katz. 2008. *The Race Between Education and Technology.* Cambridge, Mass.: Harvard University Press.

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

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107, no. 1: 35–78.

Keynes, John Maynard. 1978. "Economic Possibilities for Our Grandchildren (1930)." In *The Collected Writings of John Maynard Keynes, Volume 9: Essays in Persuasion*, edited by Elizabeth Johnson and Donald Moggridge: 321–32. London: Royal Economic Society.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull, and Giovanni L. Violante. 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica* 68, no. 5: 1029–53.

Ricardo, David. 2004. "On Machinery." In *The Works and Correspondence of David Ricardo, Volume 1*: *On the Principles of Political Economy and Taxation*, edited by Piero Sraffa with collaboration of Maurice Herbert Dobb: 385–97. Indianapolis, Ind.: Liberty Fund.

GENERAL DISCUSSION John Haltiwanger brought up the concept of expertise, as often emphasized by David Autor.[1] He asked if the authors' model accounts for whether nonautomated tasks require workers to have expertise. Gerald Cohen relatedly questioned where expertise—or, as he defined it, the need to make a judgment in the face of uncertainty (e.g., unclear instructions)—enters the authors' model, as such nonroutine tasks seem more difficult to replace with technology.

Haltiwanger further stressed that patents are an imperfect measure of innovation, considering the time needed for technology diffusion. While patent data may track technological change more closely in high-tech industries, in others (e.g., retail trade), they could overlook the complex relationship between new technologies and their effects on the labor market, Haltiwanger suggested. Barcode, for example, was patented in 1952 but did not immediately affect retail workers, and it took several decades before this technology revolutionized the sector, he noted.[2]

Katharine Abraham also emphasized the lagged relationship between innovation and labor market impact, pointing to factory electrification as another example. Moreover, in addition to time for diffusion, Abraham observed that implementing new technologies often requires large investments by firms. She asked whether the authors think artificial intelligence (AI) would follow a similar trajectory and require a similar level of investment, noting that a more rapid adoption of AI could be more disruptive to affected workers than the gradual examples of technological changes. Furthermore, Abraham echoed discussant William Kerr's suggestion of stress testing some of the authors' estimates for the pre-1940 period when validation data from the *Dictionary of Occupational Titles (DOT)* and O*NET

1. See, for example, David Autor and Neil Thompson, "Expertise," *Journal of the European Economic Association* 23, no. 4 (2025): 1203–71.

2. Google Patents, "Classifying Apparatus and Method," patent 2,612,994, https://patents.google.com/patent/US2612994A/en.

are unavailable. To gauge how well the large language model (LLM) performs without validation data, she suggested the authors explicitly prompt their model to ignore *DOT* and O*NET data when assigning tasks to occupations for later years and, provided the model follows their instructions, see whether it yields similar results.

Following up to Abraham's comment on investment in AI by firms, Caroline Hoxby suggested the authors exploit the variation in firms' training data availability to refine their predictions about which occupations AI would replace.

Referring to discussant Christina Patterson's observation that the authors' model does not assign weights to tasks and has little discipline in task aggregation, Pinelopi Goldberg pointed out that the model's predictions seem to match the direction of the actual change in employment but understate the magnitude of this change. She wondered if the absence of weights causes the model to exaggerate the labor substitution effect.

Michèle Tertilt was surprised by the model's prediction that AI would reduce demand for female labor. Given that high-contact occupations—such as teachers, nurses, and childcare workers—typically held by women seem difficult to automate, Tertilt pondered if the absence of weights led to the authors' prediction. She contended that these occupations likely spend only a few hours on tasks that AI could replace, and unless one weighs more important tasks more heavily, the model may overstate the impact of automating these tasks. Janice Eberly likewise asked if the predicted fall in demand for female labor relates to the concentration of cognitive tasks in occupations typically held by women. Moreover, she inquired whether the spillover effects captured by the authors' estimates are at the occupational level or the firm level.

Tatyana Deryugina asked about the creation of new occupations, especially those made possible by innovation. She pointed to software engineers as an example of an occupation that could not exist without technological advancements and wondered if the authors' analysis, as it stands, misses these dynamics.

Adrien Auclert acknowledged that the authors' analysis could, in its present form, only present the effects of AI on relative labor demand due to the missing intercept problem. But he thought that the authors could specify the labor supply side of the model to some extent—making some assumptions (e.g., about labor supply elasticities)—and examine the aggregate effect on the labor market. David Romer suggested the authors assume flat labor supply curves to begin, as this would allow them to estimate shifts in total rather than relative labor demand.

In response to comments on the aggregation and weight assignment of tasks as well as validation of model output, Lawrence Schmidt reiterated that the LLM assigns tasks to occupations in a manner consistent with validation data from *DOT* and O*NET for the same year and its performance for several example occupations seems reasonable. Schmidt contended that the model's lack of discipline when defining tasks could implicitly generate weights, with more important tasks appearing more often in a disaggregated list of tasks. He also highlighted the value of their data's ordinal nature and the stable occupational definitions yielded by their model.

Dimitris Papanikolaou affirmed the importance of validating the model's output. To Abraham's suggestion of stress testing estimates with the LLM, he cautioned that their instructions were not always followed by the model, and that in its current form, it is difficult to know with certainty what sources the model consults. He acknowledged that their current task-based framework, which is based on the constant elasticity of substitution (CES) structure, fails to capture the differential effects observed by Patterson and other participants. He shared that they are exploring a model based on a network structure at the industry and occupation levels, which would allow certain occupations to serve as inputs into other occupations and, in turn, allow for differential spillover effects. Papanikolaou further clarified that their current framework requires one to decide how narrowly a task is defined: While running a regression or drafting a paper could constitute individual tasks, one could alternatively treat conducting research as a single task. The authors agreed that weights matter but indicated their results appear reasonable even without weights. Papanikolaou noted that they had done another research and found similar results using weights from O*NET, which suggests that the model captures the relative importance of tasks to some extent.[3] On whether LLM could be used to formulate weights, Papanikolaou and Schmidt both emphasized the limitations of LLMs, in their current form, to produce numeric outputs.

In addition to defining tasks, researchers must also delineate which tasks they expect AI to perform in their model, Papanikolaou reflected. Thus, regarding the predicted decline in female labor demand, he explained that such prediction follows from the concentration of women in clerical and administrative work, occupations whose tasks the authors believe AI will largely be able to perform. He concurred with Hoxby that access to firms'

3. Menaka Hampole, Dimitris Papanikolaou, Lawrence D. W. Schmidt, and Bryan Seegmiller, "Artificial Intelligence and the Labor Market," working paper 33509 (Cambridge, Mass.: National Bureau of Economic Research, 2025).

training data would allow them to validate their predictions and specify more accurate counterfactual scenarios.

Papanikolaou and Schmidt recognized that the creation of new occupations requires a conceptual understanding of who performs new tasks, with Papanikolaou noting that the literature has yet to determine why tasks are bundled together in certain occupations. Schmidt added that modeling the creation of new tasks is possible but measuring the emergence of these tasks—and who does them—remains difficult. In future research, the authors hoped to investigate how task descriptions have changed over time, although such an exercise seems infeasible given LLM's current limitations.

In terms of technological exposure and spillovers, Schmidt concurred that if most of a worker's tasks are exposed to technological changes, the substitution effect could be substantial; however, he emphasized, it is rare that technology can do everything this worker is doing. Schmidt suggested that the most exposed workers are likely to work in industries witnessing the most technological progress and hence these workers could also benefit from such progress the most. Thus, while their analysis concludes that higher exposure on average hurts workers, it is not the case that new technologies invariably harm exposed workers, Schmidt resolved. In fact, it appears that the positive spillovers from the first wave of AI were large enough to offset the direct displacement effects, he noted. Papanikolaou conceded that the authors' spillover estimates are somewhat attenuated, because their analysis lacks information on firms, a potentially important piece given many spillovers take place at the firm level.

Considering the effects on nonroutine tasks in particular, Schmidt posited that innovations that automate nonroutine tasks could trigger worker substitution rather than replacement outright. He discussed how, in the post-1980s period, such innovations increased labor demand while still displacing older, higher-paid incumbent workers by redistributing jobs toward younger workers. He added that variation across cohorts suggests that younger workers seem more responsive to technological exposure: Perhaps, they are more willing to develop the expertise that a technology cannot or has not yet acquired.

On the use of patent data and technology diffusion, Papanikolaou agreed that patents overlook diffusion but pointed out that other data—such as information from 10-K forms or earnings calls—are also imprecise. He expressed hope, however, that advances in LLMs will make new data available to economic historians. Schmidt further noted that the AI adoption data currently available are heavily biased to firms. He agreed that patents are an imperfect measure as many AI advancements are not patentable. He likewise voiced an interest in exploring new data sources.

To comments by Auclert and Romer on general equilibrium, Papanikolaou responded that although their model could be augmented to consider impacts on aggregate labor demand under assumptions about labor supply responses and industry-level spillovers, it would be difficult to discipline these assumptions. Jón Steinsson supported the authors' decision to set aside general equilibrium considerations and instead focus on changes in relative labor demand. Steinsson noted that balanced growth preferences—which imply that the income and substitution effects on labor supply cancel each other out—should serve as the authors' benchmark, as one would need to articulate why their model deviates from such a benchmark before arguing for aggregate impacts.

Appendix

A Model Appendix

Section A.1 contains the key equilibrium conditions characterizing the solution of our model. We then discuss how we compute the solution of our quantitative exercises related to AI in section A.2

A.1 Key Model Equations

Here, we reproduce the key equations which are necessary for solving for the equilibrium of our model. For further details, including detailed derivations, we refer the model to Hampole et al. (2025). Our model is essentially nested in theirs with two minor changes. First, we assume perfect competition in labor and product markets, which implies that the terms reflecting markups and monopsony wedges (Θ and $\mathcal{M}(j)$ in Hampole et al. (2025)) both equal one and can be omitted from the labor demand equation. We now proceed to reproduce the main equations of the model.

The industry level CES:

$$\bar{Y} = \left(\int_{\mathcal{I}} \alpha_I^{\frac{1}{\theta}} Y_I^{\frac{\theta-1}{\theta}} dI \right)^{\frac{\theta}{\theta-1}} . \tag{A.1}$$

The occupation level CES:

$$Y_I = \left(\int_{\mathcal{O}} \alpha_I(o)^{\frac{1}{\chi}} Y(o, I)^{\frac{\chi-1}{\chi}} \right)^{\frac{\chi}{\chi-1}} . \tag{A.2}$$

The task level CES:

$$Y(o, I) = \left(\sum_j \alpha_o(j)^{\frac{1}{\psi}} y(j)^{\frac{\psi-1}{\psi}} \right)^{\frac{\psi}{\psi-1}} , \tag{A.3}$$

The labor/capital level CES:

$$y(j) = \left(\gamma_j \, l(j)^{\frac{\nu-1}{\nu}} + (1-\gamma_j) \, k(j)^{\frac{\nu-1}{\nu}} \right)^{\frac{\nu}{\nu-1}} . \tag{A.4}$$

Effective Labor Supply:

$$\ell(j) = \alpha(j)^{\beta} h(j)^{1-\beta}. \tag{A.5}$$

Optimal Hours Allocation

$$h(j) = \frac{\alpha(j) w(j)^{\frac{1}{\beta}}}{\sum_{k \in J} \alpha(k) w(k)^{\frac{1}{\beta}}}. \tag{A.6}$$

Occupation Level Efficiency

$$X(o, I) = \left[\sum_{j \in J} \alpha(j) p(j)^{1-\psi} \right]^{-\frac{1}{1-\psi}} = P(o, I)^{-1} \tag{A.7}$$

Firm Level Efficiency

$$Z_I \equiv \left(\int_{\mathcal{O}} \alpha_I(o) P(o,I)^{1-\chi} \right)^{-\frac{1}{1-\chi}} = P_I^{-1} \tag{A.8}$$

Firm Inverse Demand Curve

$$Y_I = \alpha_I Z_I^{\theta} \bar{Y} \tag{A.9}$$

Marginal Cost of task

$$p(j) = \left(a_j\, w(j)^{1-\nu} + b_j\ q(j)^{1-\nu} \right)^{\frac{1}{1-\nu}} \tag{A.10}$$

Demand for occupation o output

$$Y(o,I) = \alpha_I(o) P(o,I)^{-\chi}\, Z_I^{-\chi}\, Y_I \tag{A.11}$$

Demand for task j output:

$$y(j) = \alpha_I \alpha_I(o) \alpha(j) p(j)^{-\psi} X(o,I)^{\chi-\psi} Z_I^{\theta-\chi} \bar{Y}. \tag{A.12}$$

Demand for task j labor

$$l(j) = \alpha_I \alpha_I(o) \alpha(j) \frac{a_j}{w(j)^{\nu}} \left(a_j w(j)^{1-\nu} + b_j\, q(j)^{1-\nu} \right)^{\frac{\nu-\psi}{1-\nu}} \times X(o,I)^{\chi-\psi}\, Z_I^{\theta-\chi} \bar{Y}. \tag{A.13}$$

Demand for task j capital

$$k(j) = \alpha_I \alpha_I(o) \alpha(j) \frac{b_j}{q(j)^{\nu}} \left(a_j w(j)^{1-\nu} + b_j\, q(j)^{1-\nu} \right)^{\frac{\nu-\psi}{1-\nu}} \times X(o,I)^{\chi-\psi}\, Z_I^{\theta-\chi} \bar{Y}. \tag{A.14}$$

Labor Supply for task j

$$L_o(j) = N(o,I)\, \alpha(j)^{\beta} h(j)^{1-\beta} = N(o,I)\, \alpha(j) w(j)^{\frac{1-\beta}{\beta}} \left(\sum_{k \in J} \alpha(k) w(k)^{\frac{1}{\beta}} \right)^{\beta-1} \tag{A.15}$$

Worker supply for occupation o

$$N(o,I) = \alpha_I \alpha_I(o) W(o,I)^{\zeta} = \alpha_I\, \alpha_I(o) \left(\frac{W(o,I)}{\bar{W}} \right)^{\zeta}. \tag{A.16}$$

where

$$\bar{W} = \left[\int \int \alpha_I\, \alpha_I(o) W(o,I)^{\zeta+1} dI do \right]^{\frac{1}{\zeta+1}}. \tag{A.17}$$

Which depends on total wages for occupation o

$$W(o,I) \equiv \sum_{j \in J_o} \alpha(j)^{\beta} h(j)^{1-\beta} w(j) = \left(\sum_{j \in J} \alpha(j)\, w(j)^{\frac{1}{\beta}} \right)^{\beta} \tag{A.18}$$

Labor market clearing equation: labor supply equal to labor demand.[8]

$$w(j)^{\frac{1}{\beta}} \left(\sum_{j \in J_o} \alpha(j) w(j)^{\frac{1}{\beta}} \right)^{\beta-1+\zeta\beta} \bar{W}^{-\zeta} \varphi = a_j\, w(j)^{1-\nu} \left(a_j w(j)^{1-\nu} + b_j\, q(j)^{1-\nu} \right)^{\frac{\nu-\psi}{1-\nu}} X(o,I)^{\chi-\psi} Z_I^{\theta-\chi} \bar{Y}. \tag{A.19}$$

A.2 AI counterfactual calculations

To compute the implied changes and employment and wages associated with our AI counterfactual, we use the following procedure. We assume that the economy begins in a symmetric steady state in which $w(j)$, $q(j)$, and γ_j are the same across all tasks. In this case, we can then pin down all alpha parameters using the observed allocation of labor across industries and occupations, with the total amount of labor normalized to 1. Without loss of generality, we separately identify γ_j from $q(j)$ by normalizing that $w(j) = q(j)$, then picking the common task price which is consistent with a marginal cost of per-task output of 1 and our target for the capital share. We then pin down φ to match the level of aggregate labor supply associated with this initial allocation.

Next, to solve for the new steady state, we look for a growth rate of aggregate wages and output which satisfies the equilibrium conditions specified above. Rather than rely on the loglinear approximation presented in the text, we solve the equilibrium conditions exactly. This process is greatly facilitated by the fact that equation (A.19) is the same across all tasks that have the same $\varepsilon(j)$ within the same occupation-industry cell.

B Data Construction

B.1 Obtaining technology descriptions from patent data

B.1.1 Patent data

Our patent data comes from two sources. The pre-1976 patents (1836 - 1975) are scraped from Google patents. Specifically, from year 1836 to 1975 (including), there are in total 3,924,149 patents (according to PatentMetrics). Post-1976 patents are downloaded directly from PatentsView. PatentsView contains the full text of all patent documents from 1976 to 2024. There are in total

[8] Here φ is an aggregate labor supply normalizing constant which plays no role except to ensure that initial task wages are constant and equal to capital prices in symmetric initial equilibrium for the model simulation, which we discuss in appendix D.

8,206,179 patents that have a numeric patent number (we focus on these patents). We construct patent text data by combining the patent title, abstract, brief summary, detailed description, and claim (use the order as described).

B.1.2 Extracting technology descriptions from patent documents

Our goal is to use the textual description of these innovations to relate them to the tasks that workers do. Unfortunately, however, the quality of the extracted text of patent documents varies considerably across years. Therefore, we use a modern large language model (LLM) to obtain a concise description of the key new innovation from each patent document. In particular, we feed each patent document into *llama-3.1-8B-Instruct*, along with the following prompt:

> Here is a description of an invention from a patent issued in [PATENT ISSUE YEAR HERE]. Summarize the following: 1. what the invention does; 2. what is the main innovation; 3. how it can be used in production at that time; 4. which industry or industries are most likely to make use of this invention. Give detailed answers to each question, but each answer should be no more than four sentences long. Please avoid extrapolating too much. Put your answer to the four questions within a tuple as (1 &&&&& [ANSWER 1] 2 &&&&& [ANSWER 2] 3 &&&&& [ANSWER 3] 4 &&&&& [ANSWER 4]). In other words, use &&&&& to split answers to different questions, and replace [ANSWER #] with the appropriate answer to each question. All output should be contained within the tuple (do not output any other things). [PATENT TEXT HERE].

The large language model we use has a maximum input capacity of approximately 131,000 tokens. For patents exceeding this limit, we truncate the content and retain only the first 100,000 tokens. These longer patent texts are more common in patents issued after 2000. If the LLM returns an empty summary, we use *llama-3.1-70B-Instruct* to re-summarize those problematic summaries after the first round. We are able to do so for all except for 3,141 out of all 3,924,149 pre-1976 patents. Among them, there are 1215 patents that have empty patent text that can not be scrapped, 3 patents that have 404 website error. Among all 8,206,179 post-1976 patents, we are able to extract valid summary for all except for 30,477 (0.37%) patents.

We then obtain a vector representation of these summaries using *openai: text-embeddings-3-small* model. Thus, each patent can be represented as a vector of length 1,536. Note that summaries include the brief summary of invention ([Answer 1] in the query), the main innovation of the patent ([Answer 2] in the query), the production use of the patent ([Answer 3] in the query). We use the industry of use of the patent ([Answer 4] in the query) to assign industry code.

B.1.3 Patents to Industries

We use the same LLM model (*llama-3.1-8B-Instruct*) to obtain a coarse measure of which industries are particularly likely to see productivity improvements from technology p. As we note above, our LLM prompt includes a request for providing a likely industry of use. We use this to obtain a coarse measure of which industries are particularly likely to see productivity improvements from technology p. We obtain the vector representation of the industry-of-use descriptions using the same embedding model and compute the cosine similarities with the textual descriptions of the ind1950 industries from Census.

We assign a patent as relevant to a given Census industry if the Census industry title is the most textually similar to the patent's industry of use description (based on cosine similarity of the embeddings generated for the patent industry of use and Census industry title); we further assign the patent to the industry if the textual similarity is in the top 1% of all industry-patent similarity pairs for patents issued in that decade. In particular, for each calendar year, we determine a similarity threshold by taking the 99th percentile of scores between 5,000 randomly sampled patents' embeddings and all ind1950 entries (or, if fewer than 5,000 patents exist in a given year, the entire patent set). We then calculate the pairwise similarity between every patent embedding and each industry embedding: an industry is assigned to a patent if it is either the most similar industry overall or if its similarity exceeds the year-specific 99th-percentile threshold. Because the Census "ind1950" classification is available for all Census years in our sample, we rely on this industry classification scheme throughout.

B.1.4 Clustering Patents

To categorize technological innovations over time, we cluster patents into 20 distinct categories each time period using vector representations derived from the patent summaries. Recognizing that the nature of technological advancements evolves across historical periods, we split the sample into three eras: 1850–1920, 1920–1980, and 1980–present. For each period, we apply the mini-batch K-means algorithm to cluster all patents within that timeframe. To assign meaningful labels to each cluster, we sample 500 patents per cluster and use GPT-4o to generate descriptive group names based on the sampled content.

B.2 Tasks

B.2.1 Creating Time-varying Task Data

Next, we obtain a time varying description of the key tasks performed by each occupation in each decade in our sample. In particular, for each Census occupation in each decade from 1850 to 2010, we submit the following query to *gpt-4o-search-preview*

> You are a labor economist and an expert on economic history studying the evolution of work. Describe the list of tasks that the occupation [occupation title] would perform in their day to day job in [year]. Be very careful to make your answer specific to that time period. Your answer should be a list of tasks in JSON format, and should be written in the style of occupation task descriptions in the O*NET database. Please be sure to avoid duplicating the same task within your list; each task should be conceptually distinct. At the end please briefly provide a comprehensive and highly specific list of your sources as well. [Formatting instructions follow]

We constrain the search context to the United States and configure the search context size to its maximum setting. For each decade–occupation pair, the LLM generates a set of task titles and corresponding descriptions. We concatenate each title with its description to form the final task text representation. The result of this query is a combination of 84,393 task–occupation–decades, spanning the years 1850 to 2010 (in 10 year intervals, excluding year 1890 as the Census microdata for that year are not available). On average, the search query returns approximately 14 tasks per occupation in each decade. We then represent these tasks as numerical vectors using the *openai: text-embeddings-3-small* embeddings from OpenAI. For robustness, we perform this exercise using two alternative LLMs: gpt-4o-2024-11-20 (referred to as gpt-4o) and Meta-Llama-3.1-405B-Instruct (Llama). Our baseline is gpt-4o-search-preview, due to its integrated web search capabilities.

These task descriptions obtained from the LLM will play a key role in our empirical analysis. To ensure that these descriptions are indeed reliable, we perform a validation exercise using the occupation task descriptions from the Dictionary of Occupational Titles (DOT) and O*NET. In particular, we compare the semantic meaning of all these task descriptions between the LLM and the DOT or O*NET counterparts, whenever these are available. Thus, for each occupation in the 1940, 1980, and 2010 decades, we compute the average task embedding of the LLM task descriptions and compute its cosine similarity with the average task embedding of its DOT or O*NET counterpart. We then compare the distribution of these similarity scores to a placebo distribution where we randomly compare the LLM tasks with a different occupation in the DOT or O*NET. As we see in Appendix Figure A.1, the task descriptions we obtain from the LLM are fairly similar to their DOT or O*NET counterparts from the same period, with the within-occupation cosine similarity between LLM tasks and DOT/O*NET having a distribution that is substantially shifted to the right relative to the cross-occupation placebo. Interestingly enough, the LLM tasks from the 2010 period are particularly close to the O*NET tasks, which suggests that these tasks were part of the LLM's search query.

B.2.2 Task Classification - Manual, Cognitive, Interpersonal

We further use the same LLM we used to extract our occupation task description to classify tasks into three categories: manual, cognitive, and interpersonal. To do so, we first obtain a complete list of broad work activities from O*NET (edition 28.3). O*NET work activities are a less specific categorization of the content of work than the detailed occupation tasks. There are 41 distinct work activities, which we partition into manual, cognitive, and interpersonal categories. Our approach is similar to Acemoglu and Autor (2011), who also inspect and group O*NET work activities into different task types in order to broadly understand the task content of occupations. We list these work activities falling under a given category to give textual context for the large language model when classifying a given task. We submit the following prompt to *gpt-4o-search-preview*

> You are an expert labor economist. Your job is to read descriptions of a task performed by labor and classify whether the task is most likely a "manual task", "interpersonal task" or "cognitive task". A manual task would be expected to emphasize at least some of the following characteristics: Performing General Physical Activities; Handling and Moving Objects; Controlling Machines and Processes; Operating Vehicles, Mechanized Devices, or Equipment; Repairing and Maintaining Mechanical Equipment; Repairing and Maintaining Electronic Equipment; Inspecting Equipment, Structures, or Materials; Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment. A interpersonal task would be expected to emphasize at least some of the following characteristics: Communicating with Supervisors, Peers, or Subordinates; Communicating with People Outside the Organization; Establishing and Maintaining Interpersonal Relationships; Assisting and Caring for Others; Selling or Influencing Others; Resolving Conflicts and Negotiating with Others; Performing for or Working Directly with the Public; Coordinating the Work and Activities of Others; Developing and Building Teams; Training and Teaching Others; Guiding, Directing, and Motivating Subordinates; Coaching and Developing Others; Providing Consultation and Advice to Others; Staffing Organizational Units. A cognitive task would be expected to emphasize at least some of the following characteristics: Getting Information; Monitoring Processes, Materials, or Surroundings; Identifying Objects, Actions, and Events; Estimating the Quantifiable Characteristics of Products, Events, or Information; Judging the Qualities of Objects, Services, or People; Processing Information; Evaluating Information to Determine Compliance with Standards; Analyzing Data or Information; Making Decisions and Solving Problems; Thinking Creatively; Updating and Using Relevant Knowledge; Developing Objectives and Strategies; Scheduling Work and Activities; Organizing, Planning, and Prioritizing Work; Documenting/Recording Information; Interpreting the Meaning of Information for Others; Working with Computers; Performing Administrative Activities; Monitoring and Controlling Resources. #################### Here is the task text: [TASK TEXT HERE] Based on the above definition, would this be considered a manual task, interpersonal task, or cognitive task? Give a manual/interpersonal/cognitive/uncertain answer, along with a brief (no more than one sentence) explanation. Your answer should be in the form of a tuple (answer, explanation). Only include this tuple in your response.

Among all the 84,393 task–occupation–year triplets, 45% are classified as manual, 34% as cognitive, and 21% as interpersonal.

B.2.3 Task Classification - Specific Vocational Preparation

For our AI simulations in section 4 of the main text, we assume that cognitive tasks are exposed to AI. We classify cognitive tasks in O*NET using the exact prompt listed above. However, we also assume that high-expertise cognitive tasks are not exposed to AI. We measure expertise by mapping it the concept of "Specific Vocational Preparation" as defined by O*NET (see `https://www.onetonline.org/help/online/svp`.). In particular, we use a prompt originally constructed by Kogan et al. (2023)–who demonstrate that the categories generated from this prompt are strongly predictive of an occupation's O*NET's job zone category, which captures occupational skill and training required. We send the following prompt to the OpenAI GPT-4o Search model:

> 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. You must format the answer in a tuple '(A/B/C/D, one-sentence reasoning)'. The occupation task is: $<$ Insert Task Description $>$

We classify 4.9% of O*NET tasks into the more than 5 year category; 32.4% into the 1 to 5 year category; 39.3% into 3 months to 1 year; and 23.4% to fewer than 3 months.

B.3 Employment

We obtain employment counts using the IPUMS Census extracts.

B.3.1 IPUMS Population Sample

We use the full population sample for the period 1850–1940, and use the 1% sample for 1950, 5% sample for 1960, 1% state Form 1 sample for 1970, 5% state sample for 1980 and 1990. Due to the smaller sample sizes for the ACS, in 2010, we use 2008-2012 ACS 5-year sample and take all data for years from 2018-2012, as an average centered at year 2010. Similarly for 2020, we use 2018-2022 ACS 5-year sample, as an average centered at year 2020.

We use USA IPUMS census data to extract information related to population, including gender, age, occupation, industry, labor force participation, education, weeks/hours worked, and income:

Decade(s)	Data Source and Sample Description
1850–1940	Full population sample
1950	1% sample
1960	5% sample
1970	1% state Form 1 sample

1980	5% state sample
1990	5% state sample
2010	2012 ACS 5-year sample
2020	2022 ACS 5-year sample

We restrict the sample to census respondents aged 15 to 75 that report that they are employed. (census variable labforce = 2 for years 1850–1920, and empstat = 1 for other years). We exclude members of the armed forces and occupation codes indicating a non-occupational response, such as helping at home (occ1950 codes 975–999). Following Katz and Margo (2014), we include both men and women in our analysis; however, we also investigate the robustness of our findings to restricting to men. We compute employment in a specific occupation (or industry, i.e. ind1950 or in2d, when available) using the Census responded weights (census variable perwt).

B.3.2 Consistent Occupation Codes

To compute employment growth across decades, we aggregate the decade–specific definition of census occupations into time-consistent classifications between the two start- and end-of-period Census years involved in the employment change. For occupation classifications, we rely on the 1950 Census occupation coding scheme (IPUMs variable "occ1950") for employment change observations with start-of-period Census years before 1980. We use a revision of the Census 1990 scheme created by Autor et al. (2024) (called "occ1990dd_18") for employment change observations with start-of-period years 1980 and later. Because there is less resolution of "occ1950" codes available in IPUMs data in the year 1940 or post-1980, we convert 1950 occupations into a collapsed scheme (also introduced by Autor et al. (2024) and called "occ1950rj") for any employment changes where the start-of-period or end-of-period includes 1940, or if the end-of-period of the employment change is post-1980.

The detailed table with our time-consistent occupation coding scheme is below:

Employment StartYear	Employment EndYear	OccScheme
1850	1860	occ1950
1850	1870	occ1950
1860	1870	occ1950
1860	1880	occ1950
1870	1880	occ1950
1880	1900	occ1950
1900	1910	occ1950
1900	1920	occ1950
1910	1920	occ1950
1910	1930	occ1950
1920	1930	occ1950
1920	1940	occ1950
1930	1940	occ1950rj
1930	1950	occ1950
1940	1950	occ1950rj
1940	1960	occ1950rj
1950	1960	occ1950
1950	1970	occ1950
1960	1970	occ1950
1960	1980	occ1950rj
1970	1980	occ1950rj
1970	1990	occ1950rj

1980	1990	occ1990
1980	2000	occ1990
1990	2000	occ1990
1990	2010	occ1990
2000	2010	occ1990
2000	2020	occ1990

We use matching weights $w_{o' \to o}$ to construct a weighted crosswalk from the year-specific occupation code o' to the time-consistent code o we use above. We take $w_{o' \to o}$ to be the share of individuals in a time-consistent occupation o who are also assigned to the year-specific occupation o' in the given Census decade based off Census employment counts. As we explain in the next section, we use these to constructed weighted versions of time-consistent occupational technology exposure mean and concentration across tasks.

C Measurement

Here, we discuss our measurement approach in detail.

C.1 Task exposure

Let T denote a decade period (e.g. 1940-49, 1950-59); j, an occupation task; p, a patent, and similarity$_{p,j}$ the cosine similarity between the OpenAI text embeddings for patent p and task j. Our goal is to determine whether a given patent exposes a particular task. We impose sparsity in this relationship, by assuming that only pairs sufficiently in the right tail of the distribution of similarity$_{p,j}$ are related. For our baseline we impose a 95th percentile cutoff. Denote this 95th percentile cutoff in the distribution of task-patent similarity scores as $P95$.

Due to trends in language, we de-mean similarity scores by year and compute percentile cutoffs separately for the tasks in each Census decade. Specifically, for a task j which is relevant to a Census occupation in decade T, we compute the $P95$ percentile cutoff of the similarity between all patents issued in the 20 years before and after the start of Census decade T and decade T tasks. For example, for tasks in Census year 1940, the $P95$ threshold is calculated from the similarity score distribution between all patents issued in 1920   1959 and all decade 1940 tasks.

Since it will be computationally costly to get the percentile, we take a 1/10000 sample from all patent - task pairs except for decade 2000. For decade 2000, there are around 10,000 tasks and 3 million patents (1980 - 2019) in the window, which forms a similarity sequence of length 30 billion. We take a 1/1000000 sample from it to compute $P95$.

We compute two versions of technological exposure of task j in decade T, mapped to either a 20-year or 10-year forward-looking horizon ($H = 10$ or $H = 20$). Let P_H denote the set of patents issued in the time window $[T, T + H)$ (so when $T = 1940$, P_{10} includes patents issued from 1940-1949, inclusive, and P_{20} includes patents issued from 1940-1959, inclusive) and N_H the number of patents in P_H. The exposure of the task j to technology is simply the probability it is relevant to patents issued over that horizon:

$$\text{Exposure}_{j,T}^{H} = \frac{1}{N_H} \sum_{p \in P_H} \mathbf{1}(\text{similarity}_{p,j} > p95) \tag{A.20}$$

C.2 Constructing Occupation-Level Measures of Exposure

The model labor demand equation (10) implies that the means and variances of task-level exposure are a sufficient statistic for occupational employment growth. We therefore take the means and variances (which is our measure of exposure concentration) of task-level exposure across all tasks j which are relevant to Census occupation o. This yields our measures of occupational average task exposure to technology, and Exposure Concentration in occupational exposure to technology. Specifically, let $J(o, T)$ denote the set of tasks j relevant to occupation o in Census decade beginning in T. We have

$$\text{Mean Exposure}_{o,T}^{H} = \frac{1}{|J(o,T)|} \sum_{j \in J(o,T)} \text{Exposure}_{j,T}^{\text{H}} \tag{A.21}$$

and

$$\text{Exposure Concentration}_{o,T}^{H} = \frac{1}{|J(o,T)|} \sum_{j \in J(o,T)} \left(\text{Exposure}_{j,T}^{\text{H}} - \text{Mean Exposure}_{o,T}^{H}\right)^2 \tag{A.22}$$

for $H = 10, 20$.

To use the measure to examine employment across time, we map year-specific occupation-level exposure to time-consistent occupation schemes, following the procedure discussed in Section B.3.2.

There is one nuance to consider for occupation-level exposures. While we observe and predict employment-growth using the cross-time consistent schemes discussed previously in the appendix, we generate task exposures for year-specific titles corresponding to time-varying Census occupation coding schemes. In practice, the year-specific Census occupation codes for which we construct LLM-generated occupational tasks may map to more than one of the time-consistent codes in the list of time-consistent occupation codes used for each year. In such a case, we weight tasks for a year-specific code according to its employment weight within the time-consistent code in that year: if a year-specific occupation o' has 5 tasks in decade T, and it makes up 10% of overall employment in a time-consistent occupation o in decade T, then a single task j in occupation o' will receive a $10\% \times 1/5 = 0.5\%$ weight in computing occupation o overall mean exposure in decade T. Similarly, the same weight is applied when calculating exposure concentration across all tasks.

More formally, if a time-specific Census occupation o' has $J_{o'}$ tasks, in a given decade, and a share $w_{o' \to o}$ in time-consistent code o come from year-specific census code o', then each task j in occupation o' receives weight $\omega_j \equiv (1/J_{o'}) \times w_{o' \to o}$ in computing the occupation mean and concentration in equations (A.21) and (A.22). That is, in practice we calculate

$$\text{Mean Exposure}_{o,T}^{H} = \sum_{j \in J(o,T)} \omega_j \times \text{Exposure}_{j,T}^{\text{H}} \tag{A.23}$$

and

$$\text{Exposure Concentration}_{o,T}^{H} = \sum_{j \in J(o,T)} \omega_j \times \left(\text{Exposure}_{j,T}^{\text{H}} - \text{Mean Exposure}_{o,T}^{H}\right)^2. \tag{A.24}$$

C.3 Task Type Decomposition of Occupational Exposure - Manual, Cognitive, and Interpersonal

As discussed above in Section B.2.2, we classify each task to one of the three types: manual, cognitive and interpersonal. We use these categories to compute the technology exposure to the manual, cognitive or the interpersonal part of the occupation. The measure construction is almost the same as above, with the difference of using only tasks that are of speicific type for any occupation. Specifically, denote task type $\tau \in \{\text{manual}, \text{cognitive}, \text{interpersonal}\}$, Mean Exposure$_{o,T}^{\tau,H}$ the type τ exposure of occupation o in decade T, for H horizon. We use $J(o,T,C)$ for tasks relevant to occupation o in decade T which is classified as type τ.

$$\text{Mean Exposure}_{o,T}^{\tau,H} = \frac{1}{|J(o,T,\tau)|} \sum_{j \in J(o,T,\tau)} \text{Exposure}_{j,T}^{\text{H}} \tag{A.25}$$

We also define the type τ task share of occupation o in decade T, as the share of tasks classified as type τ for all tasks occupation o performs in decade T:

$$\text{Task Share}_{o,T}^{\tau} = \frac{|J(o,T,\tau)|}{|J(o,T)|} \tag{A.26}$$

Using the definition above, we see that the occupational mean exposure is equal to the sum of the exposure of the manual, cognitive and interpersonal mean task exposure, multiplied by their respective task shares:

$$\text{Mean Exposure}_{o,T}^{H} = \sum_{\tau} \text{Task Share}_{o,T}^{\tau} \times \text{Mean Exposure}_{o,T}^{\tau,H} \tag{A.27}$$

C.3.1 Productivity Spillover

We compute the industry spillover measure using the growth rates in the number of patents issued to each industry. The industry assignment for each patent is according to section B.1.3.

Specifically, for H-year employment changes, our industry spillover measure is

$$\text{Spill}_{I,T}^{H} = \log(\text{Matched Patents}_{T \to T+H,I}) - \log(\text{Matched Patents}_{T-H \to T,i}) \tag{A.28}$$

where Matched Patents$_{T \to T+H,i}$ is the number of patents issued within window $[T, T+H)$ that are matched to industry I.

C.4 Constructing a shift-share IV

Our shift-share identification strategy builds on Acemoglu et al. (2016), and leverages how the arrival of breakthrough technologies diffuses to 'downstream' labor-saving technologies. The construction of the shift-share instrument entails two steps. The first step (the shift) involves predicting the arrival of breakthrough patents in a given tech class c at time t based on innovation in other tech classes in the past,

$$\lambda_{c,t,\tau} = \sum_{c' \neq c} \Omega_{c' \to c,t,\tau} \times I_{c',t-\tau}. \tag{A.29}$$

Here, $\Omega_{c' \to c,t,\tau}$ is a technology diffusion matrix constructed based on the textual similarity of patents: its elements are the average similarity of patents in technology class c to patents in technology class c' for patents issued in tech class c at time t and tech class c' at time $t-\tau$; hence τ represents a diffusion lag period for innovation to propagate from tech class c' to c. When constructing Ω, we set its diagonal to zero so that we only use spillovers to class c from other technology classes c'. $I_{c,t}$ gives the intensity of breakthrough patents in class c and year t, as measured by the share of patents that are breakthroughs at time t and in technology class c.

We then use a Poisson model to predict the number of patents issued in year t for tech class c, where the independent variable is the average value of $\lambda_{c,t,\tau}$ across diffusion lags $\tau = 5, \ldots, 20$. We allow the coefficients in the Poisson model to vary by issue decade T; this measure is a strong predictor of subsequent patenting in tech class c, with an average t-stat of 4.2 across the decade-by-decade coefficients. Finally, we aggregate yearly predicted patenting from the Poisson model to the decade level. Finally, we divide by the total patents issued in that decade to get the tech class predicted patenting share in decade T. Call the predicted tech class number of patents $\widehat{N}_{c,T}$ and the predicted share of total patenting $\widehat{S}_{c,T}$.

To construct the second step (the shares) we estimate the likelihood that a patent j from technology class c is related to task j within the window $[T-10, T-10+H)$:

$$\alpha^H_{j,T} = \frac{1}{N^H_{c,T-10}} \sum_{p \in P^H_{c,T-10}} \mathbf{1}(\text{similarity}_{p,j} > P95). \tag{A.30}$$

where $P_{c,T-10}$ is the set of patents issued to tech class c in decade $T-10$.

Our shift-share measure for the innovation exposure of task j is:

$$Z^H_{j,T} = \sum_c \alpha^H_{j,T} \times \widehat{N}^H_{c,T}, \tag{A.31}$$

Finally, we compute the mean and variance of $Z^H_{j,T}$ across all tasks j which are applicable to occupation o. This yields our two IVs for a 10-year horizon of analysis, which we denote $Z^{H,Mean}_{o,T}$ and $Z^{H,Concentration}_{o,T}$, respectively. The IVs at the 20-year horizon are constructed in the exact same manner, except we replace $\widehat{N}_{c,T}$ with the implied tech-class patenting aggregated over a two-decade forward-looking period instead of a single decade.

C.4.1 Industry Spillover IV

We instrument for industry-level technological spillovers as follows. Let $\Gamma_{I,c,T-H}$ be the probability that a patent in tech class c came from industry I based on our textual mapping of patents to industries in time period T, within time period $[T-H,T)$. We estimate $\Gamma_{I,c,T-H}$ by the number of patents in tech class c, divided by the number of patents in both tech class c and industry I, within time period $[T-H,T)$. We instrument for the predicted matched patents in the industry at the H-year horizon by taking

$$\text{Predicted Matched Patents}_{T\to T+H,I} = \sum_c \Gamma_{I,c,T-H} \times \hat{S}^H_{c,T}. \tag{A.32}$$

Then, following the definition of our industry spillover measure (A.28), we construct the IV for spillover effect at the 20-year horizon by taking the log change:

$$Spill^{IV,H}_{I,T} = \log(\text{Predicted Matched Patents}_{T\to T+H,I}) - \log(\text{Predicted Matched Patents}_{T-H\to T,I}) \tag{A.33}$$

D Model Simulation

Here, we discuss details of the model simulation.

D.1 Simulating new product creation

Our model simulation requires an assumption on the rate of AI-induced new product creation, captured by growth in the model's industry production share parameter α_I. To map this to the data, we first assume that patenting in AI generates new products. We use Pairolero et al. (2025) to label patents as AI-related. We include any patent that Pairolero et al. (2025) tag as related to machine-learning, computer vision, natural language processing, or speech recognition as AI-related, using the strictest 93% cutoff for tagging AI patents (see their paper for details). We then take a probabilistic mapping from patent 3-digit CPC technology classes to NAICS industries created by Goldschlag et al. (2019). We use the Goldschlag et al. (2019) mapping to 6-digit NAICS 2007 industries, and then sum across all 6-digit industries within a 2-digit NAICS sector to get the share of each patent assigned to that sector. We then sum across all patents for each industry and year to get the total number of patents assigned to that industry each year.

Next, we need a mapping of new patents to product creation. For this, we follow Argente et al. (2025), who report that the rate of new product creation within a firm follows the rate of new patent issuance (relative to the stock of existing patents) with an elasticity 0.0469. We create the stock of existing patents following Argente et al. (2025) by accumulating all patents assigned to the industry over the last 20-years, with a depreciation rate of 15%. The Pairolero et al. (2025) patent database coverage extends through the end of 2023, so we then create the rate of new AI patenting for each industry I as of 2023:

$$P_I = \frac{\text{New AI Patents Issued in } 2023_I}{\text{Stock of Existing Patents as of } 2023_I}. \tag{A.34}$$

Finally, since we simulate over a horizon of 10 years, we adjust P_I by calculating the expected total flow of new AI patents from the end of 2024 to 2034. To do this, we calculate the average annual growth rate in total AI patents from 2014 through 2023, which is $g \approx 6.9\%$. Extrapolating this same growth rate going forward, we then compute

$$\begin{gathered}\text{Predicted flow of New AI Patents from 2024 to } 2034_I = \\ \text{New AI Patents Issued in } 2023_I \times ((1+g)^{T+1} - (1+g))/g,\end{gathered}$$

where we use the fact that $\sum_{t=1}^{T} K \times (1+g)^t = K \times ((1+g)^{T+1} - (1+g))/g$. We set $T = 11$, since we simulate 10 years forward using BLS data as of the end of 2024, but cannot observe AI patents after 2023. We call the resulting ratio

$$P_I^{Adjusted} \equiv \frac{\text{Predicted flow of New AI Patents from 2024 to } 2034_I}{\text{Stock of Existing Patents as of } 2023_I}. \tag{A.35}$$

Finally, taking the elasticity of new product creation to the rate of new innovation from Argente et al. (2025), we calculate the 10-year impact of AI on product innovation $\Delta \log \alpha_I = 0.0469 \times P_I^{Adjusted}$. Since Goldschlag et al. (2019) does not provide a mapping of patents to the BLS 2-digit sector 55 (management of companies and enterprises), we assign the cross-sector average of $P_I^{Adjusted}$ to that particular sector.

D.2 Other Assumptions.

Calibrating the model requires us to map the observed distribution of employment shares to the model. We focus on the distribution of labor across occupation × industry cells, which we take from the BLS OEWS. Given that we do not have information about the capital intensity of each occupation, we assume that the economy enters in an initial steady state in which all occupations have the same level of capital intensity. In such a case, we can then set the weights α_I and $\alpha(o, I)$

equal to the across industry and within industry, across occupation wage bill shares and exactly match the initial allocation of labor in the economy as of 2024 in the BLS data. To capture aggregate product innovation, in the post-shock equilibrium we no longer impose that the α_I weights sum to 1 across industries. To reduce the dimensionality of the simulation, we remove all occupation–industry pairs which make up less than 0.1% of an industry's BLS employment. We also drop a small set of occupations in the BLS data for which we cannot observe task-level exposure. Last, we assume symmetric initial task-level capital prices and that the task-level labor share is $s_l = 60\%$, which we take from the latest year of US data reported in Karabarbounis and Neiman (2013). This allows us to calibrate the remaining CES task-level income share parameter γ_j, which we assume is common across all tasks. Finally, the general equilibrium labor supply constant φ in equation (A.19) is pinned down by our assumption that all $w(j) = q(j) = c$ in the initial symmetric equilibrium, where the initial wage/capital price c is identified by the labor share s_l combined with the market clearing condition defined in (A.19).

References

Acemoglu, D., U. Akcigit, and W. R. Kerr (2016). Innovation network. *Proceedings of the National Academy of Sciences 113*(41), 11483–11488.

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.

Argente, D., S. Baslandze, D. Hanley, and S. Moreira (2025, September). Patents to Products: Product Innovation and Firm Dynamics. Working paper, Yale; Federal Reserve Bank of Atlanta; Compendium Labs; Northwestern University. Accessed via URL: `https://www.davidargente.com/_files/ugd/b05023_c3467658dd6c4af2b3cd225e3e9780cc.pdf`.

Autor, D., C. Chin, A. Salomons, and B. Seegmiller (2024, 03). New frontiers: The origins and content of new work, 1940–2018*. *The Quarterly Journal of Economics 139*(3), 1399–1465.

Goldschlag, N., T. Lybbert, and N. Zolas (2019). An 'algorithmic links with probabilities' crosswalk for uspc and cpc patent classifications with an application towards industrial technology composition. *Economics of Innovation and New Technology 28*(6), 589–609.

Hampole, M., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2025, February). Artificial intelligence and the labor market. Working Paper 33509, National Bureau of Economic Research.

Karabarbounis, L. and B. Neiman (2013, 10). The Global Decline of the Labor Share*. *The Quarterly Journal of Economics 129*(1), 61–103.

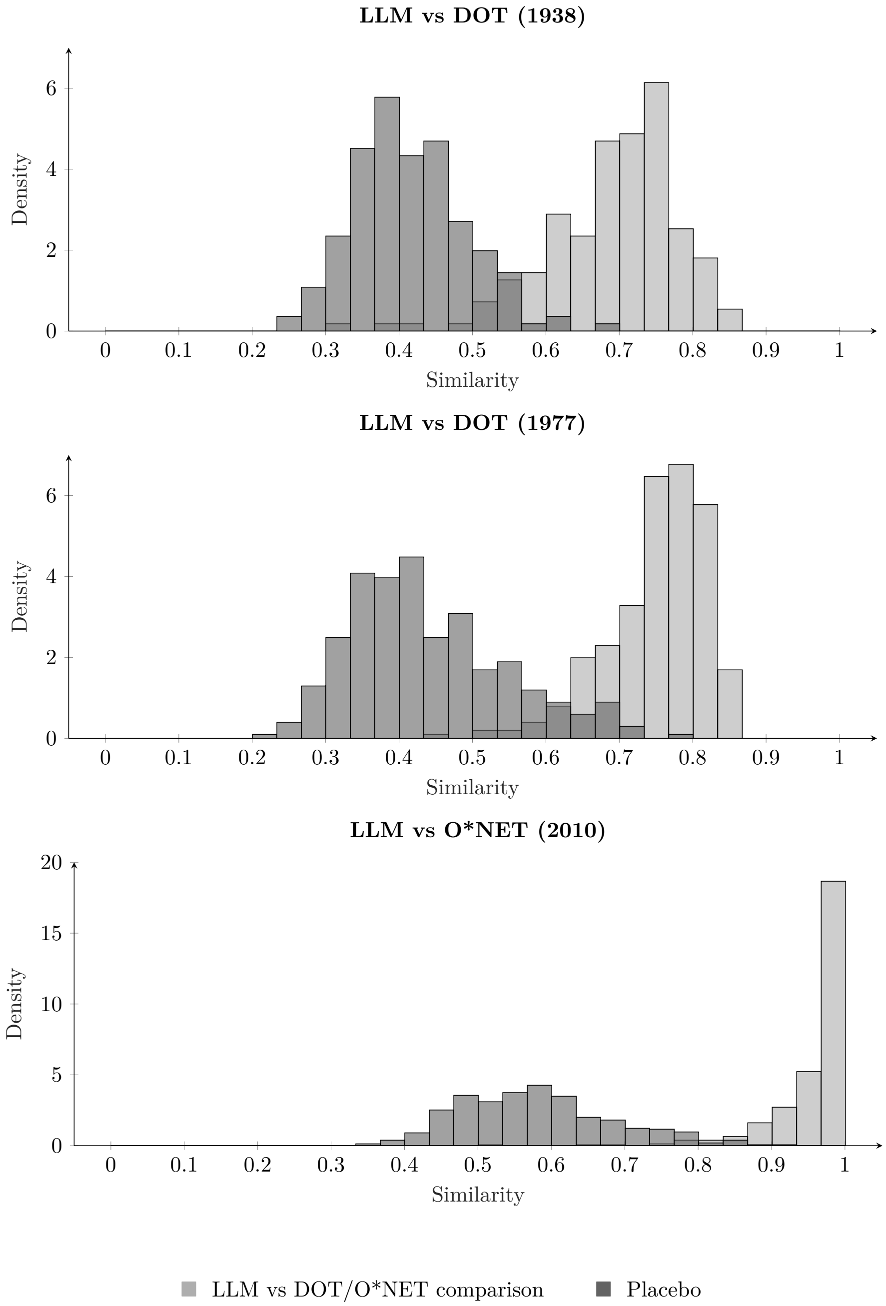
Katz, L. F. and R. A. Margo (2014, January). Technical change and the relative demand for skilled labor: The united states in historical perspective.

Kogan, L., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2023, November). Technology and labor displacement: Evidence from linking patents with worker-level data. Working Paper 31846, National Bureau of Economic Research.

Pairolero, N. A., A. V. Giczy, G. Torres, T. Islam Erana, M. A. Finlayson, and A. A. Toole (2025). The artificial intelligence patent dataset (aipd) 2023 update. *The Journal of Technology Transfer*, 1–25.

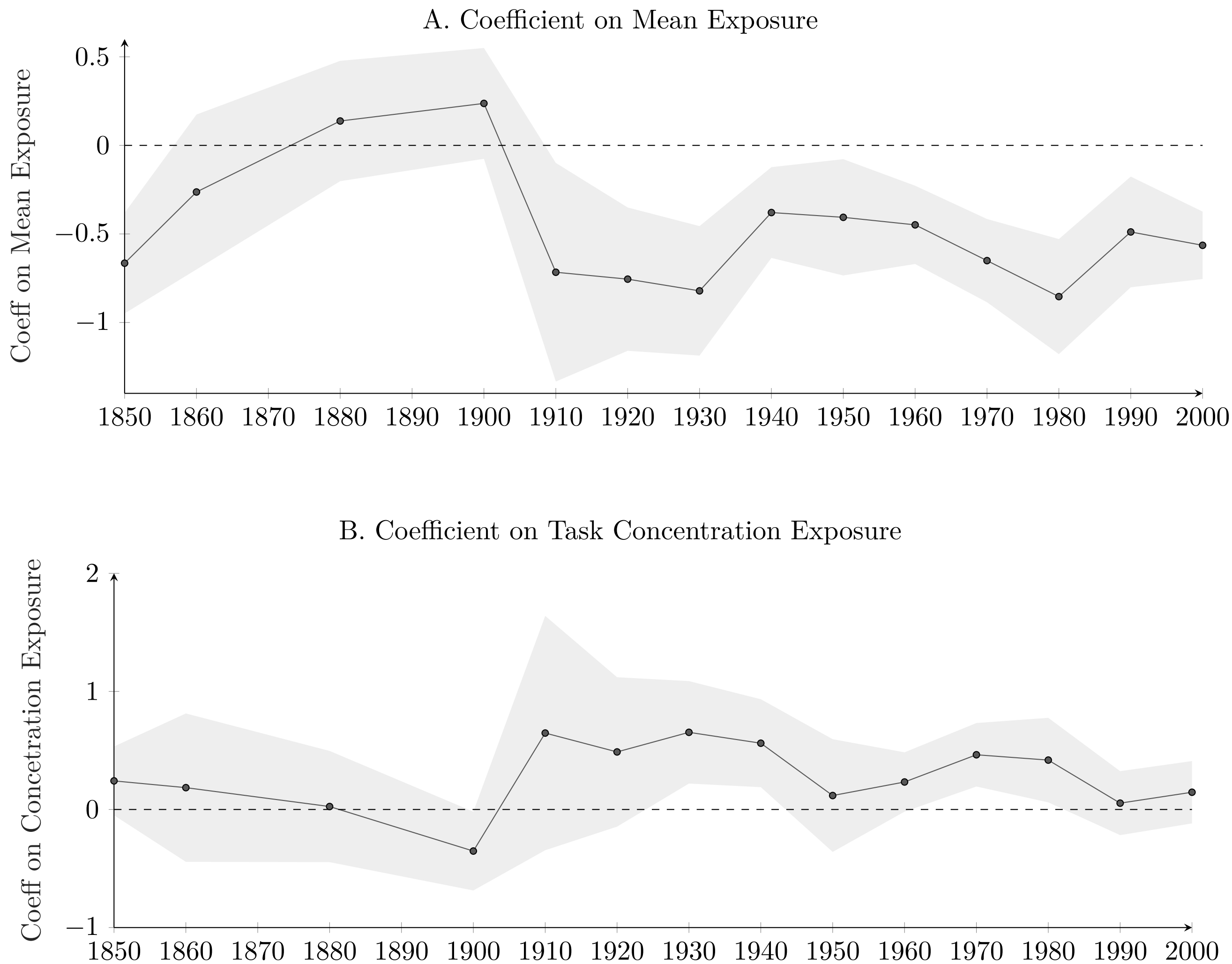
Appendix Figures and Tables

Figure A.1: Validation of LLM tasks using DOT/ONET



Note: The figure shows the distribution of cosine similarities between LLM (*gpt-4o-search-preview*) task embeddings and their DOT/O*NET counterparts. Task embeddings are averaged at the occupation level. We compare 1940 LLM tasks to 1939 DOT tasks, 1980 LLM tasks to 1977 DOT tasks, and 2010 LLM tasks to O*NET 2010 tasks. Because LLM tasks use year-specific occupation codes, whereas DOT uses occ1950rj/occ1990dd18, we match each occ1950rj/occ1990dd18 occupation to the closest year-specific occupation using the crosswalk weights. For O*NET 2010, we link SOC codes to 2010 occupation codes. A placebo distribution pairs each occupation with a random, non-matching occupation.

Figure A.2: Estimated Coefficients, by Decade

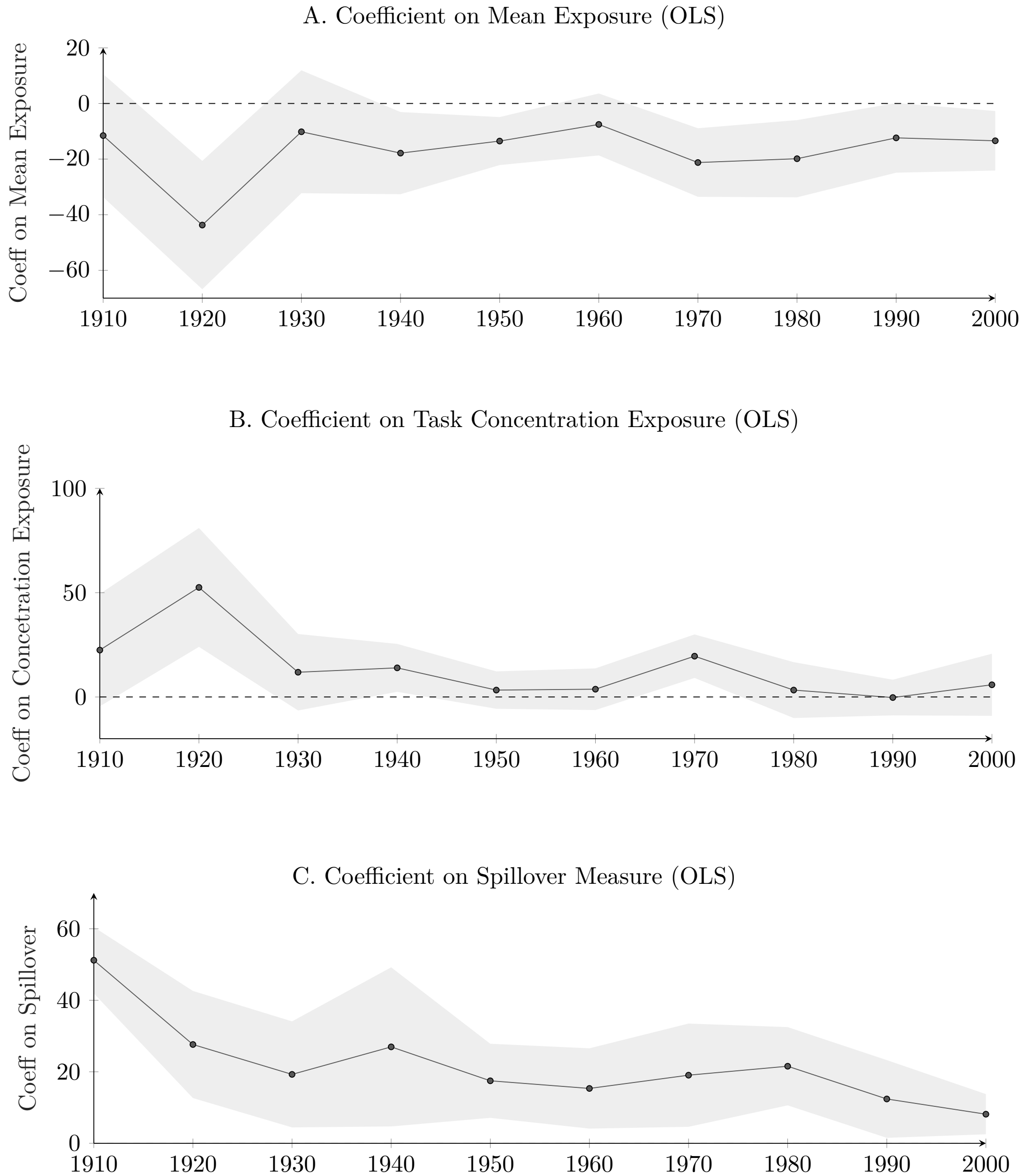


Note: This figure shows the coefficients of the following regression running for each decade T:

$$\log\left(\frac{N_{o,T+20}}{N_{o,T}}\right) = \beta\,\text{Mean Exposure}^{20}_{o,T} + \gamma\,\text{Concentrated Exposure}^{20}_{o,T} + c\,\mathbf{\Gamma}_{o,t} + \varepsilon_{o,T}.$$

Decades 1870 and 1890 are missing due to a lack of 1890 census data. Panel A shows IV coefficients of mean exposure, and Panel B shows IV coefficients of concentrated exposure. Controls $\mathbf{\Gamma}_{o,t}$ include a constant and lagged employment share $N_{o,I,T}$. Coefficients correspond to a unit standard deviation of the dependent variable and are multiplied by 100. Standard errors are clustered by occupation. Shaded bands represent 90 percent confidence intervals.

Figure A.3: Estimated Coefficients, by Decade

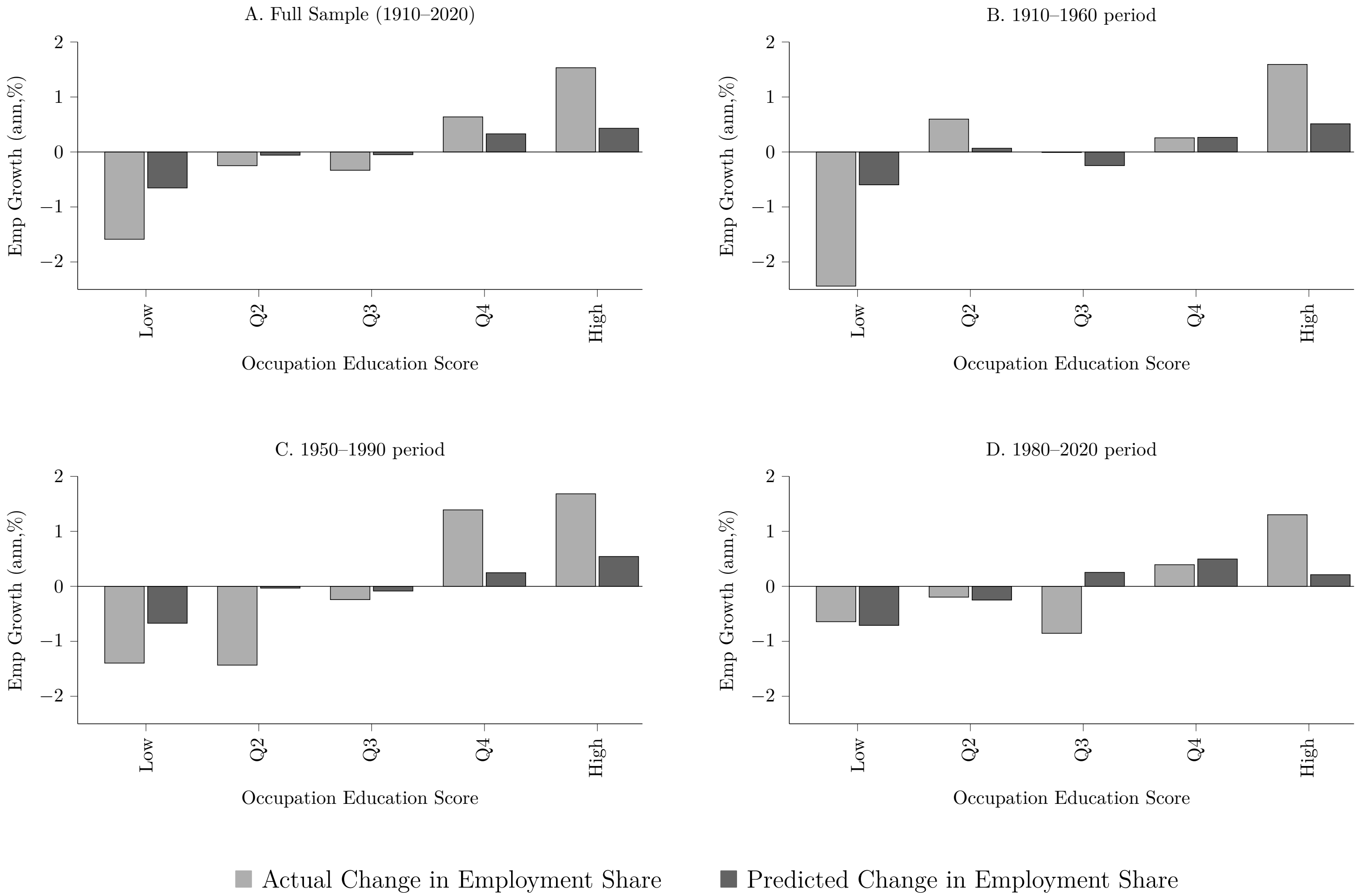


Note: Figure shows the coefficients of the following regression running for each decade T:

$$\log\left(\frac{N_{o,I,T+H}}{N_{o,I,T}}\right) = \beta\,\text{Mean Exposure}^H_{o,T} + \gamma\,\text{Exposure Concentration}^H_{o,T} + \delta\,\text{Spill}_{I,T} + c\,\mathbf{\Gamma}_{o,I,t} + \varepsilon_{o,I,T}$$

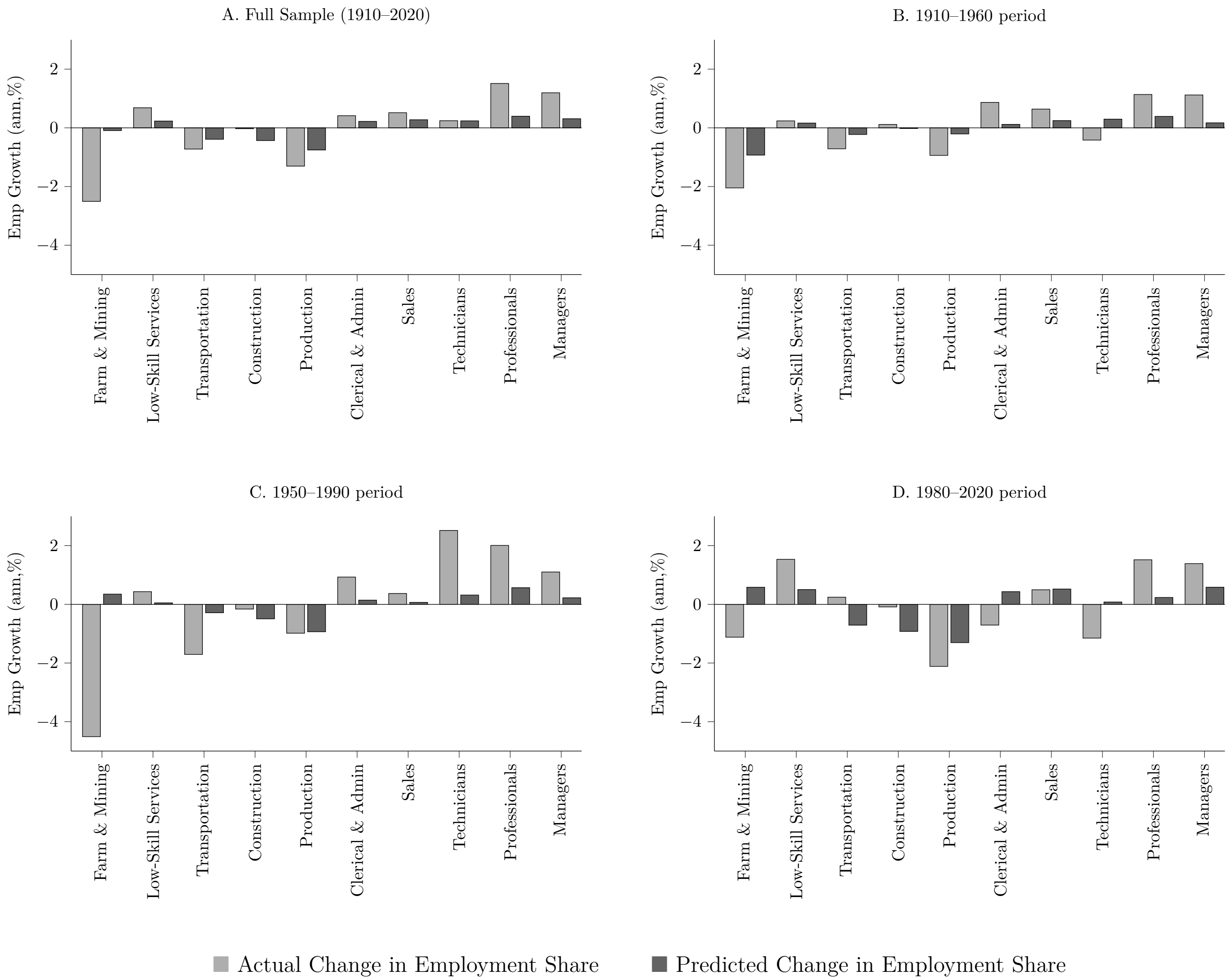
Panels A, B, and C show the coefficients of mean exposure, exposure concentration, and industry spillover, respectively. Controls $\mathbf{\Gamma}_{o,I,t}$ include sector fixed effects and lagged employment share. Coefficients correspond to a one-standard-deviation change in the dependent variable and are multiplied by 100. Standard errors are clustered by occupation and industry. Shaded bands represent 90 percent confidence intervals.

Figure A.4: Technology Exposure and Shifts in Labor Demand Across Occupations: Education



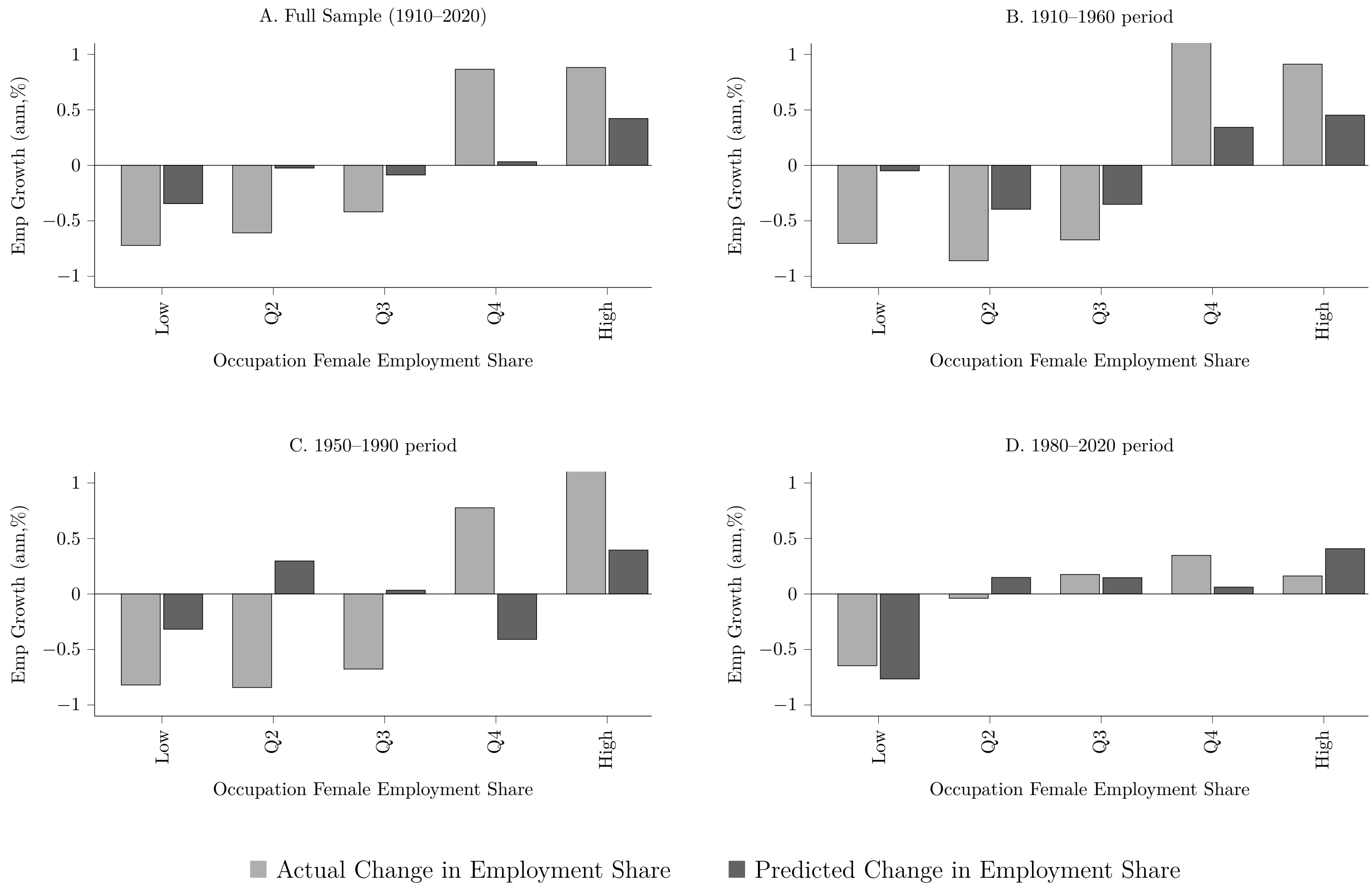
Note: This figure plots actual and technology-predicted average growth rates in employment shares based off estimates of equation (26), and also by occupational educational rank, following the procedure described in section 3.3 of the main text. In this figure we sort occupations into yearly employment-weighted quintiles based off educational attainment. We sort using the IPUMs variable “edscor50" for years before 1980 and “edscor90" for 1980 and later.

Figure A.5: Technology Exposure and Shifts in Labor Demand Across Occupations: Polarization



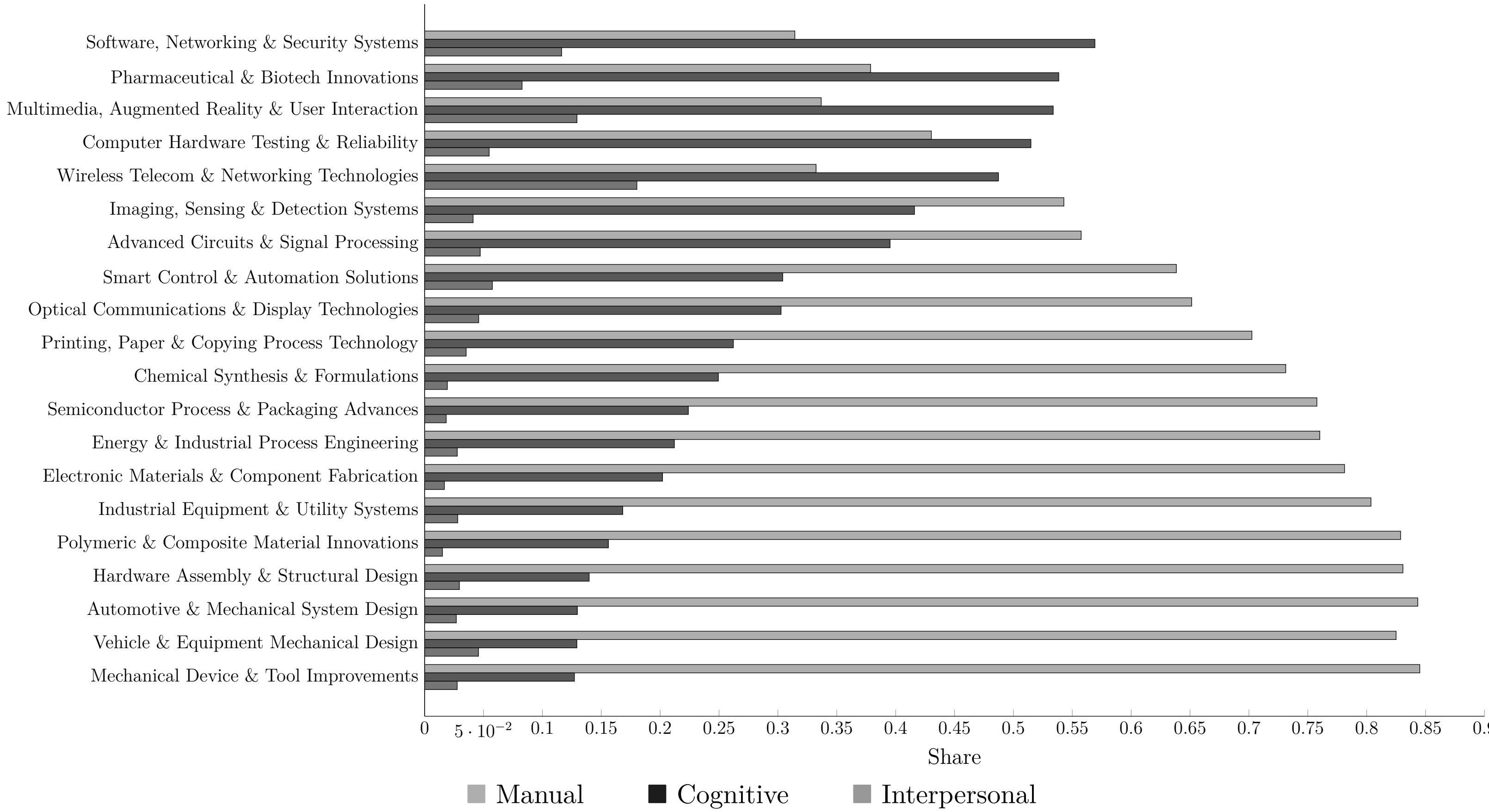
Note: This figure plots actual and technology-predicted average growth rates in employment shares based off estimates of equation (26), and also within broad occupation group categories, following the procedure described in section 3.3 of the main text. In this figure we sort occupations into broad time-consistent groups following Autor et al. (2024). Occupation groups are sorted from left to right based off their average wages.

Figure A.6: Technology Exposure and Shifts in Labor Demand Across Occupations: Female Employment Share



Note: This figure plots actual and technology-predicted average annualized growth rates in employment shares based off estimates of equation (26), and also by occupation gender composition ranking, following the procedure described in section 3.3 of the main text. In this figure we sort occupations into yearly employment-weighted quintiles based off the share of workers in the occupation who are female.

Figure A.7: Composition of Cluster Technology Exposure, by Task Type



Note: This figure plots the composition of technology exposure by each task type $\tau \in$ {Manual, Cognitive, Interpersonal} in each of the 20 clusters identified in 1980-2025 period. The composition of each type-τ task, c_τ, is defined as the share of all task exposures that are contributed by type-τ tasks in each cluster i. Specifically, we define

$$c_{i,\tau} = \frac{\sum_{T\in\mathcal{T}} \sum_{j\in K(T,\tau)} \text{Exposure}_{i,j,T}}{\sum_{T\in\mathcal{T}} \sum_{j\in K(T)} \text{Exposure}_{i,j,T}},$$

$$\text{Exposure}_{i,j,T} = \frac{1}{|P(i,T)|} \sum_{p\in P(i,T)} \mathbf{1}(\text{similarity}_{p,j} > p95)$$

where $K(T)$ is the set of all tasks in decade T, $K(T,\tau)$ contains all type τ tasks in $K(T)$, the set of decades $\mathcal{T} = \{1980, 1990, 2000\}$, and $P(i,T)$ is the set of all patents in cluster i and issued in decade T.

Table A.1: Examples of Task Sources

Occupation	**Sources**	**Tasks**
1910 blast furnaces and steel rolling mills (includes tin-plate factories)	• Xtek, Inc. - Steel Mill Jobs - A Hardworking History and a Lasting Legacy • MSU Geography - Steel Mill • Wages and Hours of Labor in the Iron and Steel Industry in the United States, 1907 to 1912 - Bulletin of the United States Bureau of Labor Statistics, No. 151 • The Working of Steel - Project Gutenberg	• Unload raw materials such as iron ore, coke, and limestone from railcars and transport them to storage areas or directly to the blast furnace charging system. • Operate skip hoists to charge the blast furnace with measured quantities of raw materials, ensuring proper layering and distribution for efficient smelting. • Monitor and control the heating of blast furnace stoves to produce the necessary hot air blast, adjusting valves and dampers to maintain optimal temperatures. • Regulate the flow of molten iron and slag from the blast furnace during tapping operations, ensuring safe and efficient discharge into ladles or molds. • Charge open-hearth furnaces with pig iron, scrap steel, and fluxes, and control the heating process to produce steel of specified composition and quality.
1940 firemen, fire department	• The Duties of a Fireman, City of Dayton Annual Report for 1940 • Job Bulletin, GovernmentJobs.com • Firefighter Job Description, City of Brentwood • Firefighter, Wikipedia	• Respond to fire alarms and extinguish fires using appropriate equipment and techniques. • Rescue individuals from burning buildings and other hazardous situations. • Operate and maintain firefighting equipment, including hoses, pumps, and ladders. • Perform routine maintenance and cleaning of fire stations, apparatus, and equipment. • Participate in regular training drills to maintain firefighting skills and physical fitness.
1950 airplane-mechanics and repairmen	• U.S. Department of Labor, Dictionary of Occupational Titles, Fourth Edition, Revised 1949 • Civil Aeronautics Administration, Aircraft Maintenance Manual, 1948 • U.S. Air Force, Technical Order 00-20A-1, Maintenance Manual, 1950	• Inspect aircraft frames, engines, and other components for wear, damage, or defects, using hand tools and visual examination. • Perform routine maintenance tasks, such as changing oil, lubricating parts, and replacing worn or defective parts. • Diagnose mechanical or hydraulic system failures to determine necessary repairs. • Repair or replace defective parts, such as wings, brakes, electrical systems, and other aircraft components. • Test aircraft systems to ensure proper functioning using diagnostic instruments.

Note: This table shows example tasks and examples of their corresponding sources as stated by *gpt-4o-search-preview*. The column "Tasks" lists examples of generated task text.

Table A.2: Technology Exposure and Employment Growth, Mean Exposure Only

	A. OLS					
	10-yr Horizon			20-yr Horizon		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean Task Exposure	-7.88***	-7.95***	-5.81***	-13.9***	-14.0***	-11.5***
	(1.17)	(1.18)	(1.21)	(2.17)	(2.17)	(2.04)
Obs	3,212	3,212	2,452	3,166	3,166	2,410
R^2 (Within)	0.015	0.022	0.042	0.021	0.032	0.062
	B. IV					
	10-yr Horizon			20-yr Horizon		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean Task Exposure	-8.39***	-8.54***	-6.42***	-15.8***	-16.1***	-13.4***
	(1.18)	(1.18)	(1.25)	(2.16)	(2.16)	(2.05)
Obs	3,212	3,212	2,452	3,166	3,166	2,410
R^2 (Within)						
F stat (Exposure)	12,586	12,604	10,846	8,485	8,489	7,194
Year FE	X	X	X	X	X	X
Employment Share, Lag		X	X		X	X
Employment Share, Lag growth			X			X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta \,\text{Mean Exposure}^{H}_{o,T} + c\,\mathbf{\Gamma}_{o,t} + \varepsilon_{o,T}.$$

for decades T spanning from 1850–2000, excluding 1890. The main variable of interest is Mean Exposure$^{H}_{o,T}$, our technology mean exposure measure, normalized to unit standard deviation. Controls $\mathbf{\Gamma}_{o,t}$ include year fixed effects, lagged employment share $N_{o,T}$, and lagged employment growth $\log\left(\frac{N_{o,T}}{N_{o,T-10}}\right)$. The controls included in each regression specification are denoted by X. Coefficients are multiplied by 100. The top panel reports the estimated coefficients using OLS, while the bottom panel reports the IV estimates. Standard errors (in parentheses) are clustered by occupation. Columns 1–3 show regressions with $H = 10$, and columns 4–6 show regressions with $H = 20$.

Table A.3: Technology Exposure and Employment Growth, Comparison across Subsamples

	A. Early Sample (1850–1920)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-6.34**	-8.79**	-4.12	2.59	-5.44**	-9.91***	-3.85	-2.95
	(2.58)	(3.47)	(4.11)	(6.17)	(2.57)	(3.50)	(4.06)	(6.71)
Concentration in Task Exposure		3.27		-8.64		6.38*		-1.22
		(3.23)		(5.61)		(3.45)		(6.60)
Obs	776	776	791	791	776	776	791	791
R^2 (Within)	0.015	0.016	0.013	0.016				
F stat (Exposure)					3,305	1,993	3,177	1,549
F stat (Concentration)					.	565	.	290
	B. Middle Sample (1910–1970)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-4.67**	-15.4***	-13.6***	-33.2***	-5.39***	-15.6***	-15.2***	-32.2***
	(2.00)	(3.68)	(3.71)	(6.27)	(2.06)	(4.29)	(3.74)	(6.82)
Concentration in Task Exposure		12.4***		23.1***		12.1**		20.5**
		(4.14)		(7.12)		(5.16)		(8.11)
Obs	1,116	1,116	1,153	1,153	1,116	1,116	1,153	1,153
R^2 (Within)	0.016	0.025	0.037	0.051				
F stat (Exposure)					6,565	3,354	8,080	4,066
F stat (Concentration)					.	898	.	956
	C. Later Sample (1960-2020)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-11.4***	-15.0***	-21.4***	-29.3***	-12.1***	-15.8***	-23.8***	-32.3***
	(1.41)	(2.71)	(2.95)	(5.00)	(1.44)	(2.80)	(2.92)	(5.21)
Concentration in Task Exposure		4.27*		9.62**		4.48*		10.7**
		(2.26)		(3.78)		(2.40)		(4.29)
Obs	1,320	1,320	1,222	1,222	1,320	1,320	1,222	1,222
R^2 (Within)	0.055	0.058	0.088	0.093				
F stat (Exposure)					14,015	7,698	5,127	2,874
F stat (Concentration)					.	4,175	.	1,790
Year FE	X	X	X	X	X	X	X	X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from the following regressions estimated for different subsamples:

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta\,\text{Mean Exposure}^H_{o,T} + \gamma\,\text{Exposure Concentration}^H_{o,T} + c\,\mathbf{\Gamma}_{o,t} + \varepsilon_{o,T}.$$

Panels A, B, and C correspond to $T \in [1850, 1900], [1910, 1950], [1960, 2000]$, respectively. Controls $\mathbf{\Gamma}_{o,t}$ include year fixed effects and lagged employment share. The coefficients correspond to a one-standard-deviation change in the dependent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation.

Table A.4: Technology Exposure and Employment Growth, Direct Effects, Employment Weights

	A. OLS					
	10-yr Horizon			20-yr Horizon		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean Task Exposure	-11.1***	-11.1***	-11.8***	-21.8***	-21.6***	-21.6***
	(2.46)	(2.58)	(2.98)	(3.89)	(4.09)	(4.22)
Concentration in Task Exposure	5.71***	5.66***	7.14***	11.2***	10.9***	12.3***
	(1.91)	(1.98)	(2.34)	(2.80)	(2.86)	(3.09)
Obs	3,212	3,212	2,452	3,166	3,166	2,410
R^2 (Within)	0.025	0.030	0.029	0.046	0.055	0.064
	B. IV					
	10-yr Horizon			20-yr Horizon		
	(1)	(2)	(3)	(4)	(5)	(6)
Mean Task Exposure	-11.1***	-11.2***	-11.9***	-23.2***	-23.4***	-22.5***
	(2.58)	(2.75)	(3.04)	(4.34)	(4.65)	(4.65)
Concentration in Task Exposure	4.46**	4.53**	5.70**	10.4***	10.4***	10.9***
	(2.13)	(2.24)	(2.51)	(3.27)	(3.42)	(3.51)
Obs	3,212	3,212	2,452	3,166	3,166	2,410
R^2 (Within)						
F stat (Exposure)	2,499	2,474	1,855	1,465	1,462	1,087
F stat (Concentration)	1,407	1,408	1,114	948	955	817
Year FE	X	X	X	X	X	X
Employment Share, Lag		X	X		X	X
Employment Share, Lag growth			X			X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta\,\text{Mean Exposure}^H_{o,T} + \gamma\,\text{Exposure Concentration}^H_{o,T} + c\,\boldsymbol{\Gamma}_{o,t} + \varepsilon_{o,T}.$$

for decades T spanning from 1850–2000, excluding 1890. The variables of interest are Mean Exposure$^H_{o,T}$, technology mean exposure, and Exposure Concentration$^H_{o,T}$, technology exposure concentration (both normalized to unit standard deviation). Controls $\boldsymbol{\Gamma}_{o,t}$ include year fixed effects, lagged employment share $N_{o,T}$, and lagged employment growth $\log\left(\frac{N_{o,T}}{N_{o,T-10}}\right)$. The controls included in each regression specification are denoted by X. Coefficients are multiplied by 100. The top panel reports the estimated coefficients using OLS, while the bottom panel reports the IV estimates. Standard errors (in parentheses) are clustered by occupation. Columns 1–3 show regressions with $H = 10$, and columns 4–6 show regressions with $H = 20$. Regressions are weighted by the employment share of which occupation; weights are right-tail winsorized at the 5% level to reduce the influence of large occupations (for example, agriculture workers).

Table A.5: Technology Exposure and Employment Growth, Subsamples, Men Only

	A. Early Sample (1850–1920)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-4.32*	-6.55*	-0.89	5.49	-3.78	-7.22**	-0.86	0.48
	(2.61)	(3.49)	(4.09)	(6.10)	(2.58)	(3.58)	(4.02)	(6.62)
Concentration in Task Exposure		2.96		-8.19		4.89		-1.83
		(3.26)		(5.50)		(3.58)		(6.46)
Obs	776	776	791	791	776	776	791	791
R^2 (Within)	0.010	0.011	0.009	0.012				
F stat (Exposure)					3,304	1,993	3,176	1,549
F stat (Concentration)					.	565	.	290
	B. Middle Sample (1910–1970)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-3.17	-12.9***	-11.1***	-33.8***	-3.96*	-13.1***	-12.7***	-33.8***
	(2.01)	(3.86)	(3.64)	(5.86)	(2.08)	(4.60)	(3.62)	(5.95)
Concentration in Task Exposure		11.2**		26.8***		10.8*		25.7***
		(4.42)		(6.03)		(5.63)		(6.07)
Obs	1,116	1,116	1,151	1,151	1,116	1,116	1,151	1,151
R^2 (Within)	0.011	0.019	0.030	0.048				
F stat (Exposure)					6,529	3,340	8,013	4,093
F stat (Concentration)					.	896	.	1,090
	C. Later Sample (1960-2020)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-9.79***	-13.2***	-18.4***	-24.1***	-10.4***	-13.8***	-20.7***	-26.8***
	(1.36)	(2.61)	(2.86)	(4.89)	(1.38)	(2.74)	(2.82)	(5.14)
Concentration in Task Exposure		4.10*		6.90*		4.15*		7.70*
		(2.20)		(3.83)		(2.38)		(4.35)
Obs	1,318	1,318	1,222	1,222	1,318	1,318	1,222	1,222
R^2 (Within)	0.044	0.046	0.070	0.073				
F stat (Exposure)					14,024	7,715	5,150	2,884
F stat (Concentration)					.	4,106	.	1,785
Year FE	X	X	X	X	X	X	X	X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from the same regressions as Table A.3, but restricts the sample to male labor only. Panels A, B, and C correspond to $T \in [1850, 1900], [1910, 1950], [1960, 2000]$, respectively. Controls $\mathbf{\Gamma}_{o,t}$ include year fixed effects and lagged employment share. Coefficients correspond to a one-standard-deviation change in the dependent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation.

Table A.6: Technology Exposure and Employment Growth, LLM Robustness: 1850–1920

	A. Baseline (GPT4o-Search)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-6.34**	-8.79**	-4.12	2.59	-5.44**	-9.91***	-3.85	-2.95
	(2.58)	(3.47)	(4.11)	(6.17)	(2.57)	(3.50)	(4.06)	(6.71)
Concentration in Task Exposure		3.27		-8.64		6.38*		-1.22
		(3.23)		(5.61)		(3.45)		(6.60)
Obs	776	776	791	791	776	776	791	791
R^2 (Within)	0.015	0.016	0.013	0.016				
F stat (Exposure)					3,305	1,993	3,177	1,549
F stat (Concentration)					.	565	.	290
	B. Alternative LLM: GPT 4o							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-1.69	-6.57	0.29	-5.43	-1.73	-6.86*	-0.94	-7.41
	(3.39)	(4.17)	(4.84)	(6.07)	(3.26)	(3.89)	(4.60)	(5.65)
Concentration in Task Exposure		4.61***		5.60***		5.06***		6.68***
		(1.45)		(2.12)		(1.38)		(1.90)
Obs	776	776	791	791	776	776	791	791
R^2 (Within)	0.009	0.014	0.011	0.015				
F stat (Exposure)					2,427	7,323	1,989	7,164
F stat (Concentration)					.	61,556	.	30,759
	C. Alternative LLM: Llama							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-1.58	1.46	1.53	5.95	-2.33	-3.53	-0.40	-1.57
	(3.60)	(6.65)	(4.86)	(8.39)	(3.39)	(5.59)	(4.71)	(7.73)
Concentration in Task Exposure		-2.92		-4.50		1.15		1.19
		(4.37)		(5.76)		(3.55)		(5.26)
Obs	776	776	791	791	776	776	791	791
R^2 (Within)	0.009	0.010	0.012	0.013				
F stat (Exposure)					1,994	1,098	1,945	1,114
F stat (Concentration)					.	1,555	.	987
Year FE	X	X	X	X	X	X	X	X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta\,\text{Mean Exposure}_{o,T}^{H} + \gamma\,\text{Exposure Concentration}_{o,T}^{H} + c\,\mathbf{\Gamma}_{o,t} + \varepsilon_{o,T}, \quad T \in [1850, 1900].$$

Panels A, B, and C correspond to measures using tasks generated by *gpt-4o-search-preview*, *gpt-4o-2024-11-20.* and *Meta-Llama-3.1-405B-Instruct*, respectively. Controls $\mathbf{\Gamma}_{o,t}$ include year fixed effects and lagged employment share. Coefficients correspond to a one-standard-deviation change in the dependent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation.

Table A.7: Technology Exposure and Employment Growth, LLM Robustness: 1910–1970 (Appendix)

	A. Baseline (GPT4o-Search)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-4.67**	-15.4***	-13.6***	-33.2***	-5.39***	-15.6***	-15.2***	-32.2***
	(2.00)	(3.68)	(3.71)	(6.27)	(2.06)	(4.29)	(3.74)	(6.82)
Concentration in Task Exposure		12.4***		23.1***		12.1**		20.5**
		(4.14)		(7.12)		(5.16)		(8.11)
Obs	1,116	1,116	1,153	1,153	1,116	1,116	1,153	1,153
R^2 (Within)	0.016	0.025	0.037	0.051				
F stat (Exposure)					6,565	3,354	8,080	4,066
F stat (Concentration)					.	898	.	956
	B. Alternative LLM: GPT 4o							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-0.96	-14.7***	-5.49	-29.4***	-2.11	-15.4***	-7.60*	-30.2***
	(2.09)	(3.35)	(4.07)	(5.57)	(2.05)	(3.36)	(3.98)	(5.52)
Concentration in Task Exposure		24.3***		41.7***		24.2***		40.8***
		(5.95)		(8.03)		(5.95)		(8.00)
Obs	1,116	1,116	1,153	1,153	1,116	1,116	1,153	1,153
R^2 (Within)	0.011	0.030	0.023	0.062				
F stat (Exposure)					5,373	3,091	7,268	5,276
F stat (Concentration)					.	1,299	.	2,659
	C. Alternative LLM: Llama							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-2.77	-7.20**	-10.7***	-21.1***	-4.11**	-9.53***	-13.0***	-25.0***
	(2.16)	(2.89)	(3.73)	(4.89)	(2.07)	(3.01)	(3.60)	(5.10)
Concentration in Task Exposure		5.83**		13.7***		7.32**		16.5***
		(2.91)		(4.64)		(3.38)		(5.19)
Obs	1,116	1,116	1,153	1,153	1,116	1,116	1,153	1,153
R^2 (Within)	0.013	0.016	0.031	0.037				
F stat (Exposure)					6,088	3,161	7,322	4,268
F stat (Concentration)					.	1,211	.	1,710
Year FE	X	X	X	X	X	X	X	X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta \,\text{Mean Exposure}^{H}_{o,T} + \gamma \,\text{Exposure Concentration}^{H}_{o,T} + c\,\mathbf{\Gamma}_{o,t} + \varepsilon_{o,T}, \quad T \in [1910, 1950].$$

Panels A, B, and C correspond to measures using tasks generated by *gpt-4o-search-preview*, *gpt-4o-2024-11-20.* and *Meta-Llama-3.1-405B-Instruct*, respectively. Controls $\mathbf{\Gamma}_{o,t}$ include year fixed effects and lagged employment share. Coefficients correspond to a one-standard-deviation change in the dependent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation.

Table A.8: Technology Exposure and Employment Growth, LLM Robustness: 1960–2020

	A. Baseline (GPT4o-Search)							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-11.4***	-15.0***	-21.4***	-29.3***	-12.1***	-15.8***	-23.8***	-32.3***
	(1.41)	(2.71)	(2.95)	(5.00)	(1.44)	(2.80)	(2.92)	(5.21)
Concentration in Task Exposure		4.27*		9.62**		4.48*		10.7**
		(2.26)		(3.78)		(2.40)		(4.29)
Obs	1,320	1,320	1,222	1,222	1,320	1,320	1,222	1,222
R^2 (Within)	0.055	0.058	0.088	0.093				
F stat (Exposure)					14,015	7,698	5,127	2,874
F stat (Concentration)					.	4,175	.	1,790
	B. Alternative LLM: GPT 4o							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-10.3***	-14.6***	-18.6***	-30.8***	-11.1***	-15.6***	-21.7***	-34.3***
	(1.35)	(2.34)	(2.82)	(4.64)	(1.38)	(2.46)	(2.84)	(4.85)
Concentration in Task Exposure		7.05**		20.3***		7.48**		21.8***
		(2.98)		(5.89)		(3.21)		(6.18)
Obs	1,320	1,320	1,222	1,222	1,320	1,320	1,222	1,222
R^2 (Within)	0.046	0.051	0.069	0.087				
F stat (Exposure)					15,201	8,043	6,250	3,329
F stat (Concentration)					.	2,683	.	1,117
	C. Alternative LLM: Llama							
	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-10.3***	-13.8***	-19.2***	-27.8***	-11.4***	-15.1***	-22.9***	-30.2***
	(1.40)	(2.14)	(2.91)	(4.04)	(1.43)	(2.22)	(2.96)	(4.39)
Concentration in Task Exposure		5.21**		12.4***		5.58**		10.8**
		(2.42)		(4.35)		(2.35)		(4.36)
Obs	1,320	1,320	1,222	1,222	1,320	1,320	1,222	1,222
R^2 (Within)	0.047	0.050	0.074	0.082				
F stat (Exposure)					11,247	5,469	4,703	2,263
F stat (Concentration)					.	2,284	.	920
Year FE	X	X	X	X	X	X	X	X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta\,\text{Mean Exposure}^{H}_{o,T} + \gamma\,\text{Exposure Concentration}^{H}_{o,T} + c\,\mathbf{\Gamma}_{o,t} + \varepsilon_{o,T}, \quad T \in [1960, 2000].$$

Panels A, B, and C correspond to measures using tasks generated by *gpt-4o-search-preview*, *gpt-4o-2024-11-20.* and *Meta-Llama-3.1-405B-Instruct*, respectively. Controls $\mathbf{\Gamma}_{o,t}$ include year fixed effects and lagged employment share. Coefficients correspond to a one-standard-deviation change in the dependent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation.

Table A.9: Technology Exposure and Employment Growth, Manual Share Control

	OLS				IV			
	10-yr Horizon		20-yr Horizon		10-yr Horizon		20-yr Horizon	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-3.55**	-3.30	-7.10***	-5.63	-4.32**	-5.86*	-7.35***	-8.64*
	(1.66)	(2.97)	(2.51)	(4.24)	(1.70)	(3.04)	(2.83)	(4.75)
Concentration in Task Exposure			3.76*	2.47			3.32	3.08
			(1.94)	(3.20)			(2.20)	(3.66)
Obs	3,212	3,166	3,212	3,166	3,212	3,166	3,212	3,166
R^2 (Within)	0.032	0.052	0.033	0.052				
F stat (Exposure)					8,251	5,875	4,474	3,038
F stat (Concentration)					.	.	1,820	1,222
Year FE	X	X	X	X	X	X	X	X
Employment Share, Lag	X	X	X	X	X	X	X	X
Share of Manual Tasks	X	X	X	X	X	X	X	X
Share of Cognitive Tasks	X	X	X	X	X	X	X	X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,T+H}}{N_{o,T}}\right) = \beta\,\text{Mean Exposure}^{H}_{o,T} + \gamma\,\text{Exposure Concentration}^{H}_{o,T} + c\,\boldsymbol{\Gamma}_{o,t} + \varepsilon_{o,T}$$

for decades T spanning from 1850–2000, excluding 1890. The variables of interest are Mean Exposure$^{H}_{o,T}$, technology mean exposure, and Exposure Concentration$^{H}_{o,T}$, technology exposure concentration (both normalized to unit standard deviation). Controls $\boldsymbol{\Gamma}_{o,t}$ include year fixed effects, lagged employment share, and the occupational share of manual and cognitive tasks. Coefficients are multiplied by 100. Standard errors (in parentheses) are clustered by occupation.

Table A.10: Technology Exposure and Employment Growth, Controlling for Industry Trends

	A. OLS							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-4.21***	-4.27***	-8.12***	-8.37***	-7.21***	-7.31***	-14.9***	-15.3***
	(1.24)	(1.24)	(2.06)	(2.07)	(2.27)	(2.32)	(3.41)	(3.48)
Concentration in Task Exposure			4.75**	4.97***			9.42***	9.82***
			(1.88)	(1.87)			(2.80)	(2.83)
N	149,217	149,217	149,217	149,217	138,923	138,923	138,923	138,923
R^2 (Within)	0.002	0.002	0.002	0.003	0.003	0.003	0.005	0.006
	B. IV							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-5.05***	-5.07***	-8.11***	-8.43***	-9.64***	-9.65***	-16.1***	-16.5***
	(1.20)	(1.21)	(2.08)	(2.09)	(2.18)	(2.21)	(3.26)	(3.32)
Concentration in Task Exposure			3.81*	4.17**			8.20***	8.78***
			(1.93)	(1.92)			(2.86)	(2.89)
N	149,217	149,217	149,217	149,217	138,923	138,923	138,923	138,923
R^2 (Within)								
F stat (Exposure)	8,398	8,482	4,390	4,438	4,677	4,745	2,473	2,511
F stat (Concentration)	.	.	1,624	1,636	.	.	1,150	1,162
Year FE	X		X		X		X	
Industry FE	X		X		X		X	
Year × Industry FE		X		X		X		X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,I,T+H}}{N_{o,I,T}}\right) = \beta \,\text{Mean Exposure}^H_{o,T} + \gamma \,\text{Exposure Concentration}^H_{o,T} + c\,\boldsymbol{\Gamma}_{o,I,t} + \varepsilon_{o,I,T}.$$

for decades T spanning from 1910–2000. The variables of interest are Mean Exposure$^H_{o,T}$, technology mean exposure, and Exposure Concentration$^H_{o,T}$, technology exposure concentration (both normalized to unit standard deviation). Controls $\boldsymbol{\Gamma}_{o,I,t}$ include year fixed effects, industry fixed effects (or year × industry fixed effects), and lagged employment share. Specific control specifications are denoted by X. Coefficients are multiplied by 100. Standard errors (in parentheses) are clustered by occupation and industry.

Table A.11: Technology Exposure and Employment Growth, Industry Controls–Innovation Spillovers, Male Only

	A. OLS							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-2.48*	-2.43*	-5.92***	-5.97***	-4.38*	-4.30*	-11.5***	-11.5***
	(1.28)	(1.28)	(2.08)	(2.07)	(2.34)	(2.34)	(3.39)	(3.41)
Concentration in Task Exposure			4.20**	4.33**			8.79***	8.85***
			(1.84)	(1.83)			(2.68)	(2.69)
Industry Spillover	8.18***	10.8***	8.19***	10.8***	17.6***	21.0***	17.6***	21.0***
	(2.46)	(1.74)	(2.46)	(1.74)	(3.38)	(3.43)	(3.38)	(3.43)
N	125,351	125,351	125,351	125,351	115,057	115,057	115,057	115,057
R^2 (Within)	0.004	0.006	0.005	0.006	0.011	0.012	0.013	0.014
	B. IV							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-2.98**	-3.16**	-5.95***	-6.17***	-6.84***	-6.73***	-12.8***	-12.9***
	(1.26)	(1.25)	(2.08)	(2.08)	(2.23)	(2.23)	(3.22)	(3.23)
Concentration in Task Exposure			3.72**	3.77**			7.70***	7.88***
			(1.85)	(1.86)			(2.73)	(2.73)
Industry Spillover	43.9**	42.2***	43.9**	42.2***	15.8**	20.0***	15.8**	20.0***
	(19.30)	(15.72)	(19.29)	(15.71)	(7.87)	(7.33)	(7.87)	(7.33)
N	124,000	124,000	124,000	124,000	114,614	114,614	114,614	114,614
R^2 (Within)								
F stat (Exposure)	4,132	4,163	2,886	2,913	2,292	2,360	1,619	1,676
F stat (Concentration)			1,056	1,355			742	761
F stat (Spillover)	7	3	6	39	27	21	18	67
Year FE	X		X		X		X	
Sector FE	X		X		X		X	
Year × Sector FE		X		X		X		X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from the same regressions as Table 2, but restricts the sample to male labor only. Controls $\mathbf{\Gamma}_{o,t}$ include year fixed effects, broad-sector fixed effects (or year × sector fixed effects), and lagged employment share. Specific control specifications are denoted by X. Coefficients correspond to a one-standard-deviation change in the independent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation and industry. Panel A reports the estimated coefficients using OLS, while Panel B reports the IV estimates. Columns 1, 2, 5, and 6 show estimation excluding exposure concentration measure, while the others include it. Columns Columns 1–4 show regressions with $H = 10$, while the others show regressions with $H = 20$.

Table A.12: Technology Exposure and Employment Growth, Industry Controls–Innovation Spillovers (1910-1970 period)

	A. OLS							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-1.78	-1.89	-9.85***	-10.0***	-3.20	-3.25	-16.2***	-16.3***
	(2.33)	(2.33)	(3.43)	(3.44)	(3.86)	(3.86)	(5.57)	(5.59)
Concentration in Task Exposure			9.91***	9.99***			16.2***	16.3***
			(2.90)	(2.91)			(4.50)	(4.54)
Industry Spillover	11.9**	16.8***	11.9**	16.8***	29.3***	32.9***	29.2***	32.8***
	(4.66)	(2.95)	(4.66)	(2.95)	(5.51)	(5.05)	(5.51)	(5.05)
N	61,669	61,669	61,669	61,669	54,821	54,821	54,821	54,821
R^2 (Within)	0.006	0.010	0.008	0.012	0.015	0.018	0.019	0.022
	B. IV							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-2.38	-2.37	-10.7***	-10.8***	-4.26	-4.32	-16.4***	-16.4***
	(2.40)	(2.38)	(3.60)	(3.60)	(3.83)	(3.83)	(5.45)	(5.49)
Concentration in Task Exposure			10.5***	10.6***			15.7***	15.6***
			(3.30)	(3.32)			(4.86)	(4.89)
Industry Spillover	86.4	53.7*	86.5	53.9*	17.7	25.2**	17.5	25.1**
	(77.92)	(28.24)	(78.03)	(28.31)	(13.30)	(12.44)	(13.35)	(12.45)
N	61,319	61,319	61,319	61,319	54,682	54,682	54,682	54,682
R^2 (Within)								
F stat (Exposure)	2,362	2,358	1,919	1,929	2,204	2,593	1,873	2,034
F stat (Concentration)			477	478			540	503
F stat (Spillover)	1	2	1	2	15	20	11	15
Year FE	X		X		X		X	
Sector FE	X		X		X		X	
Year × Sector FE		X		X		X		X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,I,T+H}}{N_{o,I,T}}\right) = \beta\,\text{Mean Exposure}^H_{o,T} + \gamma\,\text{Exposure Concentration}^H_{o,T} + \delta\,\text{Spill}_{I,T} + c\,\mathbf{\Gamma}_{o,I,t} + \varepsilon_{o,I,T}, \quad T \in [1910, 1950]$$

for decades T spanning 1910–1950. Controls $\mathbf{\Gamma}_{o,I,t}$ include year fixed effects, broad-sector fixed effects (or year × sector fixed effects), and lagged employment share. Specific control specifications are denoted by X. Coefficients correspond to a one-standard-deviation change in the independent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation and industry. Panel A reports the estimated coefficients using OLS, while Panel B reports the IV estimates. Columns 1, 2, 5, and 6 show estimates excluding exposure concentration measure, while the others include it. Columns 1–4 show regressions with $H = 10$, while the others show regressions with $H = 20$.

Table A.13: Technology Exposure and Employment Growth, Industry Controls–Innovation Spillovers (1960-2020 period)

	A. OLS							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-5.99***	-6.03***	-7.17***	-7.23***	-10.5***	-10.5***	-14.8***	-14.8***
	(1.31)	(1.31)	(2.29)	(2.28)	(2.42)	(2.42)	(4.03)	(4.04)
Concentration in Task Exposure			1.41	1.45			5.16	5.20
			(1.94)	(1.93)			(3.55)	(3.54)
Industry Spillover	4.07**	5.43***	4.07**	5.43***	10.5***	13.5***	10.5***	13.5***
	(1.86)	(2.02)	(1.86)	(2.02)	(3.55)	(3.94)	(3.55)	(3.94)
N	75,523	75,523	75,523	75,523	71,642	71,642	71,642	71,642
R^2 (Within)	0.005	0.006	0.005	0.006	0.012	0.014	0.013	0.015
	B. IV							
Employment Growth (%)	10 Years				20 Years			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean Task Exposure	-6.69***	-6.74***	-7.98***	-8.07***	-13.0***	-13.0***	-17.7***	-17.8***
	(1.29)	(1.28)	(2.18)	(2.18)	(2.36)	(2.36)	(3.80)	(3.81)
Concentration in Task Exposure			1.59	1.64			5.91*	5.96*
			(1.91)	(1.92)			(3.44)	(3.43)
Industry Spillover	36.7	36.1*	36.7	36.1*	15.6*	16.1*	15.6*	16.1*
	(23.62)	(20.80)	(23.61)	(20.79)	(9.17)	(8.17)	(9.17)	(8.17)
N	74,318	74,318	74,318	74,318	71,274	71,274	71,274	71,274
R^2 (Within)								
F stat (Exposure)	5,287	5,495	4,073	4,181	2,212	2,241	1,711	1,748
F stat (Concentration)			2,750	2,496			1,037	1,021
F stat (Spillover)	2	2	2	2	10	10	7	6
Year FE	X		X		X		X	
Sector FE	X		X		X		X	
Year × Sector FE		X		X		X		X
Employment Share, Lag	X	X	X	X	X	X	X	X

Note: The table above reports results from regressions of the form

$$\log\left(\frac{N_{o,I,T+H}}{N_{o,I,T}}\right) = \beta\,\text{Mean Exposure}^{H}_{o,T} + \gamma\,\text{Exposure Concentration}^{H}_{o,T} + \delta\,\text{Spill}_{I,T} + c\,\mathbf{\Gamma}_{o,I,t} + \varepsilon_{o,I,T}, \quad T \in [1960, 2000]$$

for decades T spanning 1910–2000. Controls $\mathbf{\Gamma}_{o,I,t}$ include year fixed effects, broad-sector fixed effects (or year × sector fixed effects), and lagged employment share. Specific control specifications are denoted by X. Coefficients correspond to a one-standard-deviation change in the independent variable and are multiplied by 100. Standard errors (in parentheses) are clustered by occupation and industry. Panel A reports the estimated coefficients using OLS, while Panel B reports the IV estimates. Columns 1, 2, 5, and 6 show estimates excluding exposure concentration measure, while the others include it. Columns 1–4 show regressions with $H = 10$, while the others show regressions with $H = 20$.