Technology-Skill Complementarity and Labor Displacement: Evidence from Linking Two Centuries of Patents with Occupations*

Leonid Kogan
MIT Sloan School of Management and NBER

Dimitris Papanikolaou
Kellogg School of Management and NBER

Lawrence D. W. Schmidt
MIT Sloan School of Management

Bryan Seegmiller
MIT Sloan School of Management

First Version: December 2019
Current Version: October 2021

Abstract

We construct new technology indicators using textual analysis of patent documents and occupation task descriptions that span almost two centuries (1850–2010). At the industry level, improvements in technology are associated with higher labor productivity but a decline in the labor share. Exploiting variation in the extent certain technologies are related to specific occupations, we show that technological innovation has been largely associated with worse labor market outcomes—wages and employment—for incumbent workers in related occupations using a combination of public-use and confidential administrative data. Panel data on individual worker earnings reveal that less educated, older, and more highly-paid workers experience significantly greater declines in average earnings and earnings risk following related technological advances. We reconcile these facts with the standard view of technology-skill complementarity using a model that allows for skill displacement.

*We are grateful to Daron Acemoglu, David Autor, Martin Beraja and seminar participants in the meetings of the Econometric Society and the Society of Economic Dynamics for valuable discussions and feedback, and to Will Cong for generously sharing his replication code. Brice Green and Jinpu Yang provided excellent research support. The paper has been previously circulated as "Technological Change and Occupations over the Long Run."

Economists and workers alike have long worried about the employment prospects of occupations whose key tasks can be easily performed by a machine, robot, software, or some other form of capital that substitutes for labor.[1] These concerns have been exacerbated by recent breakthroughs in automation technologies (e.g., software, artificial intelligence, robotics) which have expanded the set of manual and cognitive tasks which can performed by machines and have occurred contemporaneously with an increase in income inequality and a fall in the labor share of aggregate output.[2] Yet, despite the importance of these issues, systematic evidence for technological displacement remains elusive.[3] Our goal is to fill this gap: we leverage over a century and a half of data to propose and validate new metrics of workers' exposure to technological innovation and relate them to workers' labor market outcomes, both at the aggregate as well as the individual level.

To quantify workers' exposures to technical change we measure the similarity between the textual description of the tasks performed by an occupation and that of major technological breakthroughs. We identify the latter through the textual analysis of patent networks using the methodology of Kelly, Papanikolaou, Seru, and Taddy (2020). To estimate the distance between a breakthrough innovation and workers' task descriptions, we leverage recent advances in natural language processing that allow us to compute a measure of the similarity between documents that accounts for synonyms. By exploiting the timing of patent grants we can identify the extent to which certain worker groups (occupations) are exposed to major technological breakthroughs at a given point in time.

In sum, our indices capture the extent to which specific occupations are exposed to breakthrough innovations in a given year. We emphasize that, a priori, we are agnostic on whether innovations that are similar to tasks certain occupations perform are likely to be substitutes or complements. For that, we need to examine how our indicators correlate with labor market outcomes. A key advantage of our methodology is that it relies only on document text; as such, we are able to construct time-series indices of occupation exposures that span the last two centuries. For example, our technology exposure for "molders, shapers, and casters, except metal and plastic"—an occupation category which includes glass blowers as a sub-occupation—takes a relatively high value in the

[1] Fear of technological unemployment is not new. In 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." In 1811, skilled weavers and textile workers (known as Luddites) worried that mechanizing manufacturing (and the unskilled laborers operating the new looms) would rob them of their means of income. In 1930, Keynes described this type of potential labor market risk when he said, "We are being afflicted with a new disease of technological unemployment...due to our discovery of means of economising the use of labor outrunning the pace at which we can find new uses for labor." More recently, a McKinsey report estimated that between 400 million and 800 million jobs could be lost worldwide due to robotic automation by the year 2030.

[2] For instance, one of the leading explanations for the increase in the skill premium is skill-biased technical change, whereas the decline in the labor share has been attributed to capital-embodied technical change. See Goldin and Katz (2008); Krusell, Ohanian, Ríos-Rull, and Violante (2000); Karabarbounis and Neiman (2013); Acemoglu and Restrepo (2020, 2018, 2021)

[3] Due to the difficulty of constructing broad measures of labor-displacive innovations, existing work has focused on analyzing specific instances in which the impact of a specific technology on workers can be identified (Atack, Margo, and Rhode, 2019; Feigenbaum and Gross, 2020; Akerman, Gaarder, and Mogstad, 2015; Humlum, 2019).

early 1900s because of similarity with patents such as US patent number 814,612, entitled "Method of Making glass sheets." This patent relates to a technology for making glass called the cylinder machine, which allowed glass manufacturers to replace the labor of skilled hand glass blowers in favor of a highly mechanized and capital-intensive production process.[4]

Examining our technology exposure measure, we find that, prior to 1980, innovation was consistently associated with manual physical tasks; by contrast, the innovations of the late 20th/early 21st 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.

Our analysis of technical change and labor market outcomes delivers several new findings. First, we find that workers most exposed to technological breakthroughs have experienced consistently negative labor market outcomes. Indeed, a higher rate of innovation in a given industry is associated with a decline in the labor share of output even as labor productivity rises. Comparing workers across occupations differentially exposed to technology improvements, we find that an acceleration of technical change is associated with declines in employment and wage earnings for affected workers. The negative correlation between employment and our technology exposure measure is largely consistent over time—starting from the Second Industrial Revolution of the late 19th century to the present. We find some evidence that the negative relation between our measure and employment is stronger in recessions—consistent with the literature on job polarization (Jaimovich and Siu, 2018). These negative relations are estimated using a combination of public-use Census micro-data, which are available over longer horizons, and confidential administrative data from the Census which include individual tax records starting in the early 1990s.[5]

Second, we exploit the richness of administrative data to examine how these relations vary with observable characteristics, by studying a unique panel of administrative earnings records from the US

[4] Jerome (1934) documents a dramatic transformation in the production process of the glass making industry as a result of the cylinder machine: "By 1905 many hand plants had gone out of business, wages of blowers and gatherers were reduced 40 per cent, and the new machine may be said to have achieved commercial success . . . in the quarter century following the introduction of machine blowing, the window-glass industry, one of the last strongholds of specialized handicraft skill, has undergone a technological revolution resulting in the almost complete disappearance of the hand branch of the industry and the elimination of two skilled trades and one semiskilled, and also the partial elimination of the skilled flatteners."

[5] Such a negative relation is not obvious ex-ante, since our approach could in principle identify both labor-saving as well as labor-enhancing technologies. If we are primarily interested in identifying labor-saving technologies, it is possible that our measure is diluted by mixing labor-saving and productivity-enhancing innovations. To address this possibility, we exploit recent advances in topic modeling to construct a composite predictor from patent text whose purpose is to maximize the in-sample predictability of employment declines–i.e., to identify language consistent with labor saving innovations. After conducting this analysis, we find that our baseline innovation measures have a correlation of approximately 75% with this statistical factor. Comparing the performance of our measure to this benchmark, we found that both approaches lead to quantitatively similar negative outcomes in terms of both earnings and employment. On this basis, we conclude that our methodology primarily identifies labor-displacing innovations.

Social Security Administration which are linked with information on occupation and education from the Current Population Survey. Thus, relative to the literature which has mostly studied repeated cross-sections, we are able to measure a worker's occupation prior to the development of related technologies, then estimate how her earnings evolve in future years even if she switches employers, industries, and/or occupations. This analysis allows us to study the link between innovation and subsequent worker-level earnings growth rates and to address a number of potential concerns about composition effects driving our results. Our empirical analysis leverages the granularity of our patent-occupation measures to exploit variation at the industry-occupation level, i.e., in relative differences in the rate at which firms in different industries develop new technologies which are similar to a given occupation at the same point in time.

Across specifications, we find that the labor earnings of older or less educated workers are significantly more responsive to technological innovation than the average worker. More importantly, however, we find that the highest paid workers are also relatively more exposed. Specifically, workers at the top end of the earnings distribution—relative to their peers in the same occupation and in the same industry—also experience significantly larger declines than the average workers. Specifically, a one-standard deviation increase in our exposure measure is associated with a 2.5% decline in subsequent earnings compared to a 1.3% decline for the average worker. Our results are unchanged if we further control for common shocks to labor demand at the industry and occupation levels via industry-time and occupation-time fixed effects, respectively.

The fact that earnings of highly paid workers respond more to our technology exposure measure appears at first to be at odds with the canonical model of capital-skill complementarity. Indeed, a common proxy for worker skill is past income, so the fact that more highly paid workers experience larger declines seems surprising. We argue that it is not: though (a subset of) skilled workers as a group may benefit, if technological innovation is associated with skill displacement individual workers whose prior skills become obsolete may be left behind. Thus, higher paid (skilled) workers have more to lose. Consistent with this view, we find that, for the highest-paid workers (top 5% in past earnings relative to their peers in the same occupation and industry), a one-standard deviation increase in technology exposure is associated with a 1.26 percentage point increase in the probability of falling to the bottom decile of wage growth—compared to a 0.41% percentage point increase for the average worker.

To formalize this intuition, we introduce skill displacement in the canonical model of capital-skill complementarity (see, e.g. Krusell et al., 2000). We allow individual workers to supply both skilled and unskilled labor services; the quantity of skilled labor a given worker can supply depends on her skill. Improvements in technology are associated with increased likelihood of skill loss. Thus, even though skilled workers as a group (specifically, those that retain their skill) experience higher wage earnings following improvements in technology, unlucky individual workers can be left behind. On average, top workers may experience lower earnings growth following periods of technological advances if the increase in the likelihood of displacement is sufficiently high.

The calibrated model quantitatively replicates our new facts. In the model, increases in technological innovation lead to an increase in labor productivity and the skill premium—yet the labor share of output falls. On average, exposed workers in the model experience declines in wage earnings relative to peers whose skills are not related to the new technologies, and these differences are the largest for the highest paid workers. Importantly, these patterns emerge even though technology is more complementary to skilled that unskilled labor services. Following an innovation, high income workers whose skills are not displaced benefit from two forces: 1) complementarities with the more productive technology and 2) the fact that displacement of other high skilled workers' skills makes their expertise even more scarce and thus more valuable. Our model replicates our empirical result that workers with lower earnings also are hurt by the emergence of new technologies; specifically, this result obtains not because specific skills are displaced, but rather because of an increase in the supply of workers performing unskilled tasks which lowers wages.

With the model in hand, we also conduct some simple comparative statics exercises to consider the potential implications of an acceleration of the rate of innovation in the economy, consistent with the observed increase in the arrival rate of breakthrough patents which began in 1980. We consider two potential experiments. In the first case, we increase the arrival rate of new technologies but hold fixed the rate at which workers accumulate human capital. In the latter case, we also increase the rate at which workers acquire new skills so that the overall number of efficiency units of skilled human capital stays constant. In both model scenarios, such a shift generates increases in output, declines in the labor share, and increases in the skill premium in both the short and long run, all of which are consistent with trends in recent data from the US. In the former case, income inequality increases over the medium term but declines over the longer run because the higher rate of skill displacement eventually compresses the skill distribution by enough to offset the impact of a higher skill premium. In the latter case, this equalizing force is neutralized and thus income inequality increases in both the short and long run.

In sum, we provide and validate a new measure of workers' exposure to technological change that is based on the similarity between patent documents and worker job descriptions. Overall, we document a robustly negative relation between out technology exposure measure and subsequent labor market outcomes, results which are consistent with a model with capital-skill complementarity and skill displacement.

Our work contributes to the voluminous literature seeking to understand the determinants of rising inequality and the fall in the labor share. Existing work emphasizes the complementarity between technology and certain types of worker skills (Goldin and Katz, 1998, 2008; Autor, Levy, and Murnane, 2003; Autor, Katz, and Kearney, 2006; Goos and Manning, 2007; Autor and Dorn, 2013); or the substitution between workers and capital (Krusell et al., 2000; Hornstein, Krusell, and Violante, 2005, 2007; Karabarbounis and Neiman, 2013; Acemoglu and Restrepo, 2020; Hemous and Olsen, 2021). Many models in this literature treat a worker skill as a fixed characteristic and study how demand for technologies affects differences in wages between groups with different ex ante skill levels. Our contribution is to provide a direct measure of technology exposure of specific workers and examine the extent to which advances in technology are associated with differences in their labor market outcomes. Motivated by our empirical evidence, our model allows for the possibility that gains from new technologies can displace the demand for specific expertise of workers skilled at tasks associated with older vintages, similar to a literature on vintage specificity of human capital (Chari and Hopenhayn, 1991; Violante, 2002; Deming and Noray, 2020) and models which seek to explain earnings losses from job displacement via obsolescence/loss of specific human capital (Neal, 1995; Kambourov and Manovskii, 2009; Huckfeldt, 2021; Braxton and Taska, 2020).

We are not the first to analyze the differential exposure of certain occupations to technical change. Autor and Dorn (2013); Acemoglu and Autor (2011); Autor et al. (2003) document the secular decline in occupations specializing in routine tasks, starting in the late 20th century. The key idea is that routine tasks can be easily codified into a sequence of instructions. Hence, such tasks are relatively more prone to labor-saving technological change than other more complex tasks. Despite the success that this literature has had in explaining which occupations have been exposed to technologies, and what have been the effects, it is still an open question how this exposure changes over time, which technologies relate to which types of tasks and which occupations, and whether or not technological unemployment is a robust phenomenon in other time periods. More recently, Webb (2019) also analyzes the similarity between patents and occupation task descriptions. Our work differs in both scope and aim. Webb (2019) focuses on automation and the future of work, and thus restricts attention to patents identified as being related to robots, AI, or software. As a result, the analysis in Webb (2019) is largely cross-sectional in nature as he focuses on a single technological episode—the rise of AI and robots. In contrast, we construct time-series indicators to understand the relation between innovation and employment over different technological episodes and its impact on workers with different characteristics. Further, focusing on the more recent period for which wage earnings data is available (after 1980), we show that the predictability of our measure for worker earnings is complementary to the information contained in the routine-task intensity measure of Autor and Dorn (2013), the AI and robotics occupation exposure measure of Webb (2019) and is not driven by industry-specific trends.[6]

[6]In related work, Mann and Püttmann (2018) and Dechezleprêtre, Hémous, Olsen, and Zanella (2021) use patent

A significant contribution of our work lies in its scope: we provide a measure of occupational exposure to technical change that spans the period from 1850 to 2010. An important advantage of our analysis is that it allows us to draw broad conclusions regarding the relation between technical change and worker outcomes over a long time period. Further, by constructing measures at the patent-occupation level, our approach allows us to study technological change at a highly granular level. Patents also have the advantage of being associated with specific timing (filing and approval dates) and are linkable to specific firms. To this end, our empirical analysis uses this granular information to compute measures of technological change at the industry-occupation level. Our results thus complement some earlier studies which, by narrowing their scope, are able to analyze the impact of worker earnings associated with specific technologies. For example, Atack et al. (2019) analyze how workers' task transitioned from hand to machine production in the late 19th century. Recently, Feigenbaum and Gross (2020) show that incumbent telephone operators were more likely to be in lower-paying occupations following the adoption of mechanical switching technology by AT&T. Akerman et al. (2015) and Humlum (2019) provides an in depth analyses of impacts of adoption of broadband internet and industrial robots, respectively, leveraging microdata on affected workers and firms.

Though labor income risk is not the primary focus of our study, we reach a similar conclusion as Kogan, Papanikolaou, Schmidt, and Song (2020): higher-paid workers face considerably greater risk in their labor income as a result of technological innovation. Though some of the conclusions are similar, these two papers ask different questions. Kogan et al. (2020) examine the dynamics of wage earnings in response to innovation by the workers' own firm or its competitors in the product market. Kogan et al. (2020) are interested in the extent to which profit-sharing motives transfer the risk of creative destruction from the firm owners to its workers. By contrast, we examine outcomes for all workers in the same industry, differentiated by their occupation (and its exposure to major innovations). Since our goal is to capture not only innovation by a firm but also the overall adoption of a technology in a given sector, the exact origin of these innovations are not particularly relevant.

1 Motivation: Technology, Productivity, and the Labor Share

We begin with a set of facts regarding the joint dynamics of aggregate measures of innovation, measured productivity, and the labor share that serve to motivate the remainder of our analysis. To do so, we obtain data on industry-level measures of output (value added), employment and the labor share from the NBER manufacturing database—which cover the 1958 to 2018 period.

text with different classification algorithms to identify automation patents in more recent periods, though they do not relate these patents with specific occupations performing related tasks.

Measuring the degree of technological innovation that takes place in a particular industry at a particular point in time is considerably more challenging. We do so by relying on patent data and closely follow the methodology of Kelly et al. (2020). In particular, Kelly et al. (2020) identify breakthrough innovations as those that are both novel (whose descriptions are distinct from their predecessors) and impactful (they are similar to subsequent innovations). They measure a patent's novelty as its dissimilarity with the existing patent stock at the time it was filed. In particular, they construct a measure of 'backward similarity'

$$BS_j^{\tau_b} = \sum_{i \in \mathcal{B}_{j,\tau_b}} \rho_{j,i}, \tag{1}$$

where $\rho_{i,j}$ is the pairwise cosine similarity (using TF-IDF weights) of patents i and j and $\mathcal{B}_{j,\tau_b}$ denotes the set of "prior" patents filed in the τ_b calendar years prior to j's filing. Patents with low backward similarity deviate from the state of the art and are therefore novel. Similarly, they measure a patent's impact by its 'forward similarity' as

$$FS_j^{\tau_f} = \sum_{i \in \mathcal{F}_{j,\tau_f}} \rho_{j,i}, \tag{2}$$

where $\mathcal{F}_{j,\tau_f}$ denotes the set of patents filed over the next τ_f calendar years following patent j's filing. The forward similarity measure in (2) estimates of the strength of association between the patent and future technological innovation over the next τ years.

The Kelly et al. (2020) patent-level measure combines forward and backward similarity to identify patents that are both novel and impactful,

$$q_j^{\tau} = \log FS_j^{\tau_f} - \log BS_j^{\tau_b}. \tag{3}$$

To create a time-series index, Kelly et al. (2020) remove calendar year fixed effects from (3) in order to adjust for shifts in language over time. After defining a 'breakthrough' patent as one that falls in the top 10% of the unconditional distribution of importance, they then construct a time series index as the number of breakthrough inventions granted in each year.

In brief, their measure of innovation in industry j in year t is defined as

$$\psi_{j,t} = \frac{1}{\kappa_t} \sum_{p \in \Gamma_t} \alpha_{j,p}. \tag{4}$$

To construct (4), we need to determine the set of breakthrough patents that are relevant to a given industry. We first identify the set Γ_t of 'breakthrough' patents as those that fall in the top 10%

of the unconditional distribution of patent importance (based on their ratio of 10-year forward to 5-year backward similarity). We map patents to industries based on their CPC technology class using the probabilistic mapping constructed by Goldschlag, Lybbert, and Zolas (2020). Here, $\alpha_{j,p}$ denotes the probability of breakthrough patent p being assigned to industry j. Last, we scale (4) by US population κ_t and normalize to unit standard deviation.

We then estimate the following specification

$$\log X_{j,t+k} - \log X_{j,t} = \alpha(k) + \beta(k)\psi_{j,t} + \delta(k)Z_{j,t} + \varepsilon_{j,t}, \qquad k = 1 \dots T \text{ years.} \tag{5}$$

We focus on four outcome variables: output (value-added); employment; labor productivity (value-added per worker); and the labor share. We examine horizons of up to $T = 6$ years. Controls include the lagged 5-year growth rate of the outcome variable and year fixed effects.

Figure 6 plots the estimated impulse response coefficients $\beta(k)$. Panel A illustrates that an one-standard-deviation increase in the degree of innovation $\psi_{j,t}$ in a given industry is associated with an approximately 2% increase in output over the next six years. However, this increase in output primarily reflects an increase in productivity: as we see from Panels B, the overall level of employment in the industry weakly falls. As a result, we see in Panel C that, in response to a one-standard-deviation increase in $\psi_{j,t}$, labor productivity in the industry rises sharply—by approximately 3% over the next six years. Panel D illustrates that the labor share of output in the industry falls.

In brief, an increase in our technology measure is associated with higher measured labor productivity and a decline in the labor share. This pattern strongly suggests that, on average, our technology measure $\psi_{j,t}$ captures innovations that likely act as substitutes, rather than complements to labor. That is, even though improvements in technology are associated with an increase in the measured productivity of labor, they are associated with declines in the labor share. The remainder of the paper builds on this idea and aims to provide further refinement to our technology measure by identifying which set of innovations are more likely to substitute for labor inputs. In the process of doing so, we also broaden the coverage from just manufacturing to all sectors in the economy.

2 Measuring Workers' Technology Exposure

Here, we construct a measure of technological innovation occurring at a particular point in time that is relevant for a particular set of worker tasks (occupation). Our interpretation is that this measure can be used to proxy for a worker's exposure to technological innovation. To do so, we add an additional source of cross-sectional variation to the Kelly et al. (2020) time-series measure of innovation discussed above. Specifically, we recognize that breakthrough innovations at a given point in time may be differentially related to particular occupations. To construct our measure, we rely on measuring the 'distance' between the description of the technology (from the patent document) to the description of the tasks that a given occupation performs (from the Dictionary of Occupational Titles, DOT). The remainder of this section describes our methodology.

2.1 Data and Methodology

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

We obtain text data for measuring patent/job task similarity from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. We use the patent text data parsed from the USPTO patent search website in Kelly et al. (2020), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET. We then combine all tasks for a given occupation at the 2010 O*NET 6-digit level into one occupation-level corpus. See Appendix A for further details on cleaning and preparing the text files for numerical representation.

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

$$\rho_{i,j} = \frac{V_i}{||V_i||} \cdot \frac{V_j}{||V_j||} \tag{6}$$

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

This approach is often referred to as the 'the bag-of-words' approach, and has been used successfully in many settings. For example, Kelly et al. (2020) use a variant of this approach to construct measures of patent novelty and impact based on pairwise distance measures between patent documents. Since patent documents have a structure and a legalistic vocabulary that is reasonably uniform, this approach works quite well for patent-by-patent comparisons. However, this approach is less suited for comparing patent documents to occupation task descriptions. These two sets of documents come from different sources and often use different vocabulary. If we were to use the bag-of-words approach, the resulting vectors V_i and V_j would be highly sparse with most elements equal to zero, which would bias the distance measure (6) to zero.

The root cause of the problem is that the distance measure in (6) has no way of accounting for words with similar meanings. For example, consider a set of two documents, with the first document containing the words 'dog' and 'cat' and the other containing the words 'puppy' and 'kitten'. Even though the two documents carry essentially the same meaning, the bag of words approach will conclude that they are distinct: the representation of the two documents is $V_1 = [1, 1, 0, 0]$ and $V_2 = [0, 0, 1, 1]$, which implies that the two documents are orthogonal, $\rho_{1,2} = 0$.

To overcome this challenge, we leverage recent advances in natural language processing that allow for synonyms. The main idea behind this approach is to represent each word as a dense vector. The distance between two word vectors is then related to the likelihood these words capture a similar meaning. In our approach, we use the word vectors provided by Pennington, Socher, and Manning (2014), which contains a vocabulary of 1.9 million word meanings (embeddings) represented as (300-dimensional) vectors.[7] Appendix Section A.2 contains a brief discussion of how the Pennington et al. (2014) word embeddings are constructed.

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

$$X_i = \sum_{x_k \in A_i} w_{i,k} x_k. \tag{7}$$

[7] The basis for this word vector space is arbitrary; distances between word embeddings are only well-defined in relation to one another and a different training instance of the same data would yield different word vectors but very similar pairwise distances between word vectors. The two most popular approaches are the "word2vec" method of Mikolov, Sutskever, Chen, Corrado, and Dean (2013) and the global vectors for word representation introduced by Pennington et al. (2014). These papers construct mappings from extremely sparse and high-dimensional word co-occurence counts to dense and comparatively low-dimensional vector representations of word meanings called word embeddings. Their word vectors are highly successful at capturing synonyms and word analogies (vec(king) − vec(queen) ≈ vec(man) − vec(woman) or vec(Lisbon) − vec(Portugal) ≈ vec(Madrid) − vec(Spain), for example). Thus they are well-suited for numerical representations of the "distance" between words. The word vectors provided by Pennington et al. (2014) are trained on 42 billion word tokens of web data from Common Crawl and are available at `https://nlp.stanford.edu/projects/glove/`.

A key part of the procedure consists of choosing appropriate weights $w_{i,k}$ in order to emphasize important words in the document.

In natural language processing, a common approach to emphasize terms that are most diagnostic of a document's topical content is the 'term-frequency-inverse-document-frequency' (TF-IDF). We follow the same approach: in constructing (7), we weigh each word vector by

$$w_{i,k} \equiv TF_{i,k} \times IDF_k. \tag{8}$$

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

$$TF_{i,k} = \frac{c_{i,k}}{\sum_j c_{i,j}}, \tag{9}$$

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

The inverse-document frequency is

$$IDF_k = \log\left(\frac{\text{\# of documents in sample}}{\text{\# of documents that include term } k}\right). \tag{10}$$

Thus, IDF_k measures the informativeness of term k by under-weighting common words that appear in many documents, as these are less diagnostic of the content of any individual document.

In brief, $TFIDF_{i,k}$ overweighs word vectors for terms that occur relatively frequently within a given document and underweighs terms that occur commonly across all documents. We compute the inverse-document-frequency for the set of patents and occupation tasks separately, so that patent document vectors underweight word embeddings for terms appearing in many patents and occupation vectors underweight word embeddings for job task terms that appear in the task descriptions of many other occupations.

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

$$\text{Sim}_{i,j} = \frac{X_i}{||X_i||} \cdot \frac{X_j}{||X_j||} \tag{11}$$

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

In sum, we use a combination of word embeddings and TF-IDF weights in constructing a distance metric between a patent document (which includes the abstract, claims, and the detailed description of the patented invention) and the detailed description of the tasks performed by occupations. Our methodology is conceptually related, though distinct, to the method proposed by Webb (2019), who also analyzes the similarity between a patent and O*NET job tasks.[8]

2.2 Examples

To illustrate the effectiveness of our methodology in identifying links between technology and occupation task descriptions, we consider a few representative examples, some of which are summarized in Figure 1.

A key advantage of our measure is that it is available over long periods of time, and thus allows us to study very different technologies from three distinct periods of technological change–the Second Industrial Revolution of the late 1800's, the period spanning the from 1920s to around 1940, and the information technology revolution spanning the end of the 20th and beginning of the 21st centuries. For example, consider three patents in the list of breakthrough patents identified by Kelly et al. (2020). Patent 276,146, titled "Knitting Machine", was issued in the height of the Second Industrial Revolution in 1883. The occupation that is most closely related to this patent is "Textile Knitting and Weaving Machine Setters, Operators, and Tenders"; the next most similar occupation is "Sewing Machine Hand Operators", followed by "Sewers, hand". Next consider the patent for "Metal wheel for vehicles (1,405,358), which is issued in 1922. The occupation most closely related to this patent is "Automotive Service Technicians and Mechanics", with other production and metal machine workers following. Finally, we examine a patent from a very different era and representing a very different technology. The patent, entitled "System for managing financial accounts by a priority allocation of funds among accounts," is U.S. patent number 5,911,135 and was issued in 1999. The top occupations related to this patent are Financial managers, credit analysts, loan interviewers and clerks.

We next perform the reverse exercise, where we fix a particular occupation, and list the most relevant innovations. The occupations we choose are cashiers, loan interviewers and clerks, and railroad conductors. Table A.2 lists the top five patents that are linked to each of these occupations. Examining the patent tiles, we see that each one of these patents is directly related to the work performed by the given occupation. For example, one of the top patents for cashiers is

[8] Webb (2019) focuses on similarity in verb-object pairs in the title and the abstract of patents with verb-object pairs in the job task descriptions and restricts his attention to patents identified as being related to robots, AI, or software. He uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings. By contrast, we infer document similarity by using geometric representations of word meanings (GloVe) that have been estimated directly from word co-occurence counts. Furthermore, we use not only the abstract but the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention. In addition to employing a different methodology, we also have a broader focus: we are interested in constructing time-series indices of technology exposures. As such, we compute occupation-patent distance measures for all occupations and the entire set of USPTO patents since 1836.

"Vending type machine dispensing a redeemable credit voucher upon payment interrupt" (patent 5,055,657); the top patent for loan interviewers and clerks is titled "Automatic business and financial transaction processing system" (patent number 6,289,319). And finally, for rail road conductors, titled "Automatic train control system and method" (patent 5,828,979) is the top patent. In general the patents showing up on this list represent technologies that (1) relate to the work performed by individuals in that the occupation; and (2) if adopted, appear likely to be able to change the way that an occupation performs its core work functions and/or substitute for work done by that occupation.

In sum, these examples illustrate the ability of our method in identifying technologies that are related to a particular occupation. However, it is not immediately obvious whether these technologies benefitted workers in these occupations or whether they led to the displacement of workers. As a concrete example, consider US patent number 6,289,319, titled "Automatic business and financial transaction processing system", and which as shown in Table A.1 is the most similar patent to the "Loan Interviewers and Clerks" occupation. The DOT task description indicates that a person with this occupation "calls or writes to credit bureaus, employers, and personal references to check credit and personal references." The description of this patent states that "Loan processing has traditionally been a labor-intensive business...the principal object of this invention is to provide an economical means for screening loan applications." We interpret this innovation as an example of a technology which has high potential to be labor saving because it is intended to perform the same tasks performed manually by a worker in a more efficient manner.

However, there are also many counter-examples of new technologies that improve the productivity of tasks that workers are currently performing. Our exposure measure potentially also picks up these instances. For instance consider the occupation "Database Administrators" (SOC code 151141). According to the DOT, a database administrator "coordinates physical changes to computer databases." According to our distance measure, one of the most similar patents to this occupation is US patent number 5,093,782, entitled "Real time event driven database management system." This patent indicates that it provides "a database management system which is capable of supporting processes requiring the updating and retrieval of data elements at a high rate." This is likely to make the work of database administrators more efficient and hence looks more likely to be labor productivity-enhancing for this occupation.

Most likely, some of these technologies benefitted some workers at the expense of others. To illustrate the potential for such differential effects across workers of different skill levels, we consider two examples of labor saving technologies from Jerome (1934). First, consider two key innovations in the textile weaving industry during the early 20th century, the Barber-Colman warp-tying machine (patent 1,115,399) and the drawing-in machine (patent 1,364,091). Both of these technologies benefitted skilled workers at the expense of unskilled labor. Jerome (1934) notes that, the Barber-Colman warp-tying machine "will do the work of about 15 hand operators" while "it can be run by one tender." Similarly, he notes that "It is estimated that each (drawing-in machine) machine, requiring ordinarily the attention of one operator and half the time of an assistant, replaces from 5 to 6 hand drawers-in." Both of these patents are identified as breakthrough patents by Kelly et al. (2020). In terms of related occupations, our methodology identifies various types of textile workers as being the some of the most relevant.

However, not all labor-saving technologies benefit skilled labor. For instance, consider two major innovations in the window glass industry during the late 19th century—the Colburn sheet machine (patent 840,833) and the cylinder machine (patent 814,612). Following their introduction, the manufacturing process for window glass switched from being hand-made to being entirely mechanized by 1925. The displacement of skilled workers was rapid: by 1905 many hand plants had gone out of business, wages of blowers and gatherers were reduced 40 per cent. Jerome (1934) summarizes their impact thus: "In the quarter century following the introduction of machine blowing, the window-glass industry, one of the last strongholds of specialized handicraft skill, has undergone a technological revolution resulting in the almost complete disappearance of the hand branch of the industry and the elimination of two skilled trades and one semiskilled, and also the partial elimination of the skilled flatteners. The contest for supremacy now lies between the cylinder and the sheet machine processes." Both of these patents are in the top 10% of the Kelly et al. (2020) measure. In terms of our methodology, we identify "glaziers" and "molders, shapers, and casters, except metal and plastic" as being among the most related occupations to these two patents. Specifically, the latter occupation, which corresponds SOC code 519195, has a sub-occupation called "glass blowers, molders, benders, and finishers".

These two examples illustrate that the impact of a new technology on a given worker is not ex-ante obvious. Some technologies may replace un-skilled workers, while others may displace highly specialized and skilled workers. Indeed as Jerome (1934) notes, glass workers displaced by the sheet and cylinder machines in their time were considered to be members of skilled trade. Goldin and Katz (2008, Chapter 3) provide a number of historical examples where technological advances standardized tasks formerly performed by skilled artisans so that they could be performed by unskilled workers (see also Acemoglu, Gancia, and Zilibotti, 2012, for a related theoretical treatment of such a process). Further, new technologies may also generate demand for new skills—for example, the operators of the Barber-Colman warp-tying machine—hence their long-run effects may be different than their short-run impact. Hence, a significant part of the paper focuses on examining the correlation between our technology exposures and subsequent labor market outcomes.

2.3 Identifying Variation in Technology Exposures over Time

Our analysis so far delivers a measure of similarity between a given patent and a given occupation. The next step is to construct time-series indices of technological exposure at the occupation level. The key challenge in constructing a time-varying index lies in choosing how to appropriately quantify the 'degree' of innovation that occurs at a given point in time. One possibility would be to count the number of patents; however, this approach is unlikely to be fruitful, since not all patents are equally important. Various approaches have been proposed, which essentially weight patents by their forward citations (Hall, Jaffe, and Trajtenberg, 2005); estimates of their market value (Kogan, Papanikolaou, Seru, and Stoffman, 2017); or their textual similarity to prior and subsequent patents (Kelly et al., 2020).

For our purposes, we choose the Kelly et al. (2020) approach for two reasons: first, unlike forward citations, their measure is available for the entirety of our sample; second, it is available for all patents, and not just patents issued to publicly traded firms; and third, we are primarily interested in the contribution of a patent on the technology frontier rather than their private value to their firm.

We define our time series index of exposure of occupation i to technology at time t as

$$\eta_{i,t} = \frac{1}{\kappa_t} \sum_{j \in \Gamma_t} \tilde{\rho}_{i,j} \times \mathbf{1}(\tilde{q}_{j,t} \geq \tilde{q}_{p90}). \tag{12}$$

Our time-series index (12) aggregates our patent-occupation similarity scores across all breakthrough patents issued in year t. Specifically, we sum over the occupation-patent similarity score $\tilde{\rho}_{i,j}$ across the set of patents $j \in \Gamma_t$ that are issued in year t. We restrict attention to breakthrough patents, that is, patents whose Kelly et al. (2020) ratio of importance $\tilde{q}_{j,t}$ exceeds the (unconditional) 90th percentile $\tilde{q}_{p90}$.

When computing (12), we use an adjusted occupation-patent exposure metric $\tilde{\rho}_{i,j}$. Specifically, we perform the following adjustments to our raw occupation-patent exposure $\rho_{i,j}$ from (6). First, we remove yearly fixed effects. We do so in order to account for language and structural differences in patent documents over time and technology areas.[9] Second, we impose sparsity: after removing the fixed effects we set all patent $\times$ occupation pairs to zero that are below the 80th percentile in fixed-effect adjusted similarity. Last, we scale the remaining non-zero pairs such that a patent/occupation pair at the 80th percentile of yearly adjusted similarities has a score equal to zero and the maximum adjusted score equals one.

[9] Patents have become much longer and use much more technical language over the sample period, and the OCR text recognition of very early patents is far from perfect. We also slightly modify the Kelly et al. (2020) procedure by adjusting for the interaction between year and technology fixed effects, since some tech classes tend to have a naturally higher ratio of forward to backward patent similarity.

2.4 Which occupations are more exposed?

Overall, we find that over the span of our sample, service-type occupations that specialize in person-to-person interaction scoring especially low on average exposure. In particular, Table A.3 lists the top and bottom five occupations by average exposure over the time period spanning since 1850. The most exposed occupation is titled "Inspectors, Testers, Sorters, Samplers, and Weighers". The top occupations tend to be those working in production and manufacturing type jobs, which are commonly posited to be among the type of occupations most affected by new technologies. The least exposed occupations are mental health counselors, dancers, funeral attendants, judges, and clergy, all representing service job types that are unlikely to have the nature of their work substantially changed by new technologies.

Related to this point, Autor and Dorn (2013) argue that middle skill occupations have been significantly more exposed to technological innovation that low-skill workers. Using our direct measure of technology exposure, we can verify this is indeed the case. Specifically, we can examine how the technology exposure of occupations varies by 'skill levels' as proxied by wages. We obtain information on average wages by occupation from the Current Population Survey Merged Outgoing Rotation Groups (MORG, see Appendix A.4 for more details). Given the short time dimension of the data (MORG starts in 1980), we focus on cross-sectional comparisons.

Figure 2 we plot exposures against average wage percentile ranks for the post-1980 period. Consistent with Autor and Dorn (2013), we see that the most exposed occupations tend to be found in the middle of the income distribution.

2.5 Long-run trends

We begin by documenting the types of occupation that are exposed to technological innovation, and how these exposures have shifted over time. To this end, we group each occupation into eight broad categories: service; sales and office; production, transportation, and material moving; natural resources, construction, and maintenance; management, business, and financial; healthcare practitioners; education, legal, community service, arts and media; and, computer, engineering, and science. Within each of these groups we take the average of $\eta_{i,t}$ and then scale across the eight groups each year so that the total sums to one. Figure 3 plots these shares over the entire sample (1850 to the present).

Examining Figure 3, there are two points worth noting. First, 'blue-collar' occupations, that is, those related to production and construction, have been consistently more exposed to technological progress than the others. Second, this trend has materially shifted over the recent decades, possibly due to the Information Technology (IT) revolution. Starting around the 1950s, there has been a secular increase in the relative technology exposure of 'white-collar' occupations. This rise is most visible in the increased exposure of computer, engineering, and science occupations. Sales and office occupations have also seen an increased relationship with innovation, as well as management/business occupations, though these two groups remain small in their overall exposure.

A useful way of summarizing these trends is examining the characteristics of occupations most exposed to technology at a given point in time. We first examine what kinds of tasks are performed by these occupations. Basing our analysis on Acemoglu and Autor (2011), we focus on four task categories: manual tasks (routine and physical); non-routine manual (interpersonal); routine cognitive; and non-routine cognitive.[10] Let $T_w(i)$ be an indicator function that equals 1 if occupation i has a score in the top quintile across occupations for task w; also denote by ω_i the Acemoglu and Autor (2011) employment shares for occupation i. We then construct an index $\lambda_{w,t}$ of the technological exposure of task category w as follows:

$$\lambda_{w,t} = \sum_i \eta_{i,t} \times T_w(i) \times \omega_i \tag{13}$$

Figure 4 plots our measure of technology exposure $\lambda_{w,t}$, now separated for each of these task categories. The top panel (Panel A) plots these series in levels; the bottom panel (Panel B) plots their composition. The overall time-series behavior of our measures largely mimics the series of Kelly et al. (2020)—which is not surprising, given that we use their definition of breakthrough patents. We note three major innovation waves, lasting from 1870 to 1890; 1910 to 1930; and from 1970 to the present. The first peak corresponds to the beginning of the second industrial revolution, which saw technological advances such as the telephone and electric lighting and improvements in railroads. The second peak corresponds to advances in manufacturing, particularly in plastics and chemicals, consistent with the evidence of Field (2003). The latest wave of technological progress includes revolutions in information technology.

Importantly, we see that the first two major innovation waves were primarily related to occupations performing non-interpersonal manual tasks. By contrast, cognitive tasks are significantly less exposed. However, starting from the 1970s, there is a shift in the relative exposure of occupations emphasizing cognitive tasks, especially routine cognitive tasks. As a result, in the last few decades, these occupations are almost as exposed to innovation as occupations emphasizing manual tasks. This pattern is driven by information technology revolution that has led to the modern digitalization of the workplace. Occupations that relate to these type of innovations have a distinctly different

[10] Because the routine manual and non-routine manual (physical) task scores are highly correlated and also move similarly with technological exposure, we group these two task types into one category by taking the average of the two scores. For similar reasons we take the average of non-routine cognitive (analytical) and non-routine cognitive (interpersonal) to get a non-routine cognitive score.

task profile than the most prevalent technologies of past innovation waves. That said, even in the recent period, occupations emphasizing interpersonal tasks remain the least exposed to technological change. This pattern is consistent with the findings of Autor and Dorn (2013), who show that service occupations have increased in importance at the expense of occupations heavily exposed to automation, and also Deming (2017), who documents an increased importance of social skills in the labor market.

We next separate occupations by their education requirements. Specifically, we compute the share of workers in that occupation who have either completed a 4-year college degree or have attained a high-school diploma or lower in in a given year. For this analysis we crosswalk occupations to David Dorn's revised Census occ1990 level. We impute college grad and above/high school or below occupation shares for years between Census decades by linearly interpolating between the nearest available Census years and similarly interpolate occupational employment shares $\omega_{i,t}$ between Census years. We then let $S_{s,t}(i)$ be an indicator for whether occupation i is in the top quintile of the share of workers in education category s in year t. Due to data availability, we begin our analysis in 1950. We define the education exposure index $\zeta_{s,t}$ similarly to $\lambda_{w,t}$:

$$\zeta_{s,t} = \sum_i \eta_{i,t} \times S_{s,t}(i) \times \omega_{i,t} \tag{14}$$

Figure 5 presents our results. For most of the sample, we see that occupations requiring a college degree are significantly less exposed to innovation than occupations requiring a lower level of education. However, and consistent with the discussion above, we see that this pattern is shifting in the recent decades: towards the end of the sample, the difference in technology exposure between occupations requiring a college degree with those that do not has shrunk dramatically. This is especially evident in panel B of Figure 5 where we plot the composition rather than the levels of technological exposure. It's also important to note that this pattern is not driven by the increase in the share of workers with a college degree, since we assign occupations to the high education group based on their ranking in the cross-sectional distribution of occupational college grad shares. Rather, this pattern is driven by compositional shifts in the types of technologies being introduced, with an increasing share of technologies being targeted towards the tasks performed by relatively more educated occupations.

3 Technology Shocks and Labor Market Outcomes

In Section 1 we documented that breakthrough innovations are on average associated with increases in measured productivity and declines in labor share. Now, armed with an additional source of cross-sectional variation—differences in occupation exposure to specific technologies—we can examine a more direct link between technological progress and outcomes for specific workers.

We begin by focusing on group (i.e. occupation-level) outcomes in Section 3.1. The advantage of doing so is that we can examine the relation between our measure and labor market outcomes over a long period of time. The disadvantage of doing so is that patterns in some occupation-level outcomes, specifically wages, can mask important worker-level heterogeneity within the occupation. Thus, in Section 3.2 we focus on outcomes of individual workers using administrative data on worker earnings. Focusing on individual workers allows us to more closely trace the correlation between innovation and individual outcomes, while also enabling us to condition on certain observable characteristics. The disadvantage of the Census administrative data is that it is available only since the mid-1990s.

3.1 Occupation-level Evidence from repeated cross-sections

We begin by examining the relation between innovation and subsequent growth in the employment shares and average wage earnings of exposed occupations.

Data

The availability of public Census data allows us to examine employment outcomes over a long period of time (1850 to today). The Census surveys consist of repeated cross-sectional observations. Important for our purposes they contain information on occupations, which we can link to our innovation measure $\eta_{i,t}$ constructed in equation (12). Specifically, we use the 1950 Census occupation definition for pre-1950 Census years since the more updated 1990 Census classification scheme is only available in post-1950 Census years. We make use of the 1990 Census occupation classifications for the years they are available. We then crosswalk Census occupations to the David Dorn occ1990dd classification scheme using the crosswalk files provided on his website and aggregate our measure $\eta_{i,t}$ to the occ1990dd-level by averaging across 6-digit SOC codes within an occ1990dd code. This results in a Census-year by occ1990dd panel of occupation employment shares. Census records for the year 1890 were destroyed in a fire, and so the employment growth observations for the 20-year horizon in 1870 or for the 10-year horizon in 1880 are not available. The final dataset consists of an unbalanced panel of occupation–Census year employment shares and spans the Census years from 1850 to 2010. Appendix A.4 provides additional details.

In addition, we use more recent data from the Current Population Survey Merged Outgoing Rotation Groups (MORG) which provides data on both wages and employment outcomes for the post-1980 period. We use the data to create a balanced panel of wage earnings and employment growth at the level of occupation and calendar year. We obtain the cleaned MORG extracts provided by the Center for Economic Policy Research (CEPR). In particular, we use the "wage3" variable that combines the hourly earnings for hourly workers and non-hourly workers, adjusts for top-coding using a lognormal imputation, and is constructed to match the NBER's recommendation for the most consistent hourly wage series from 1979 to the present.

Technology exposure and employment (1850–present)

We examine employment outcomes using the following specification,

$$\frac{1}{k}\Big(\log Y_{i,t+k} - \log Y_{i,t}\Big) = \alpha_0 + \alpha_t + \beta(k)\,\eta_{i,t} + \lambda Y_{i,t} + \varepsilon_{i,t}, \qquad k = 10, 20 \text{ years.} \tag{15}$$

The main dependent variable $Y_{i,t}$ is the employment share in total non-farm employment. Observations are weighted by the employment share of the given occupation and standard errors are clustered by occupation. As before, $\eta_{i,t}$ is normalized to unit standard deviation. All specifications include time fixed effects; depending on the specification, we include controls for the lagged 10-year employment growth rate.

Panel A of Table 1 presents our findings. We note that there is a strong and statistically significant negative correlation between our innovation measure η and subsequent changes in employment at the occupation level. The magnitudes are significant: a one-standard deviation increase in $\eta_{i,t}$ is associated with a 0.41% annualized decline in employment over the next 10 years and a 0.70% percent decrease in employment over the next 20 years.

We next allow the slope coefficients β to vary across Census years, focusing on horizons of $k = 20$ years. Figure 7 plots the point estimates of β for each Census year along with the 90% confidence intervals based on standard errors clustered by occupation. Examining the figure, we see that the point estimates are negative for all but the 1860 and 1940 Censuses, and are significant in 1880, 1910, 1920, 1950, 1970, 1980, and 1990. The magnitude of this correlation is also fairly stable over time, implying that occupations that are exposed to innovation have had consistently experienced employment declines over the entire 150 year period.[11]

One potential concern with these findings is that they reflect industry trends. To separate our findings from industry-level sources of variation, we next aggregate the Census data at the occupation by industry level over time. We use the 1950 Census industry designations, which are available the furthest back in time. Because Census industry codes are unreliable before 1910 we start our analysis using the data from the 1910 Census. We therefore estimate a slightly modified

[11] In 1930, John Maynard Keynes wrote: "We are being afflicted with a new disease of technological unemployment...due to our discovery of means of economising the use of labor outrunning the pace at which we can find new uses for labor." (Keynes, 1930). Indeed, the 1920–40 period corresponded with a large innovation wave that was associated significant declines in employment for occupations whose tasks were related to those innovations.

version of equation (15),

$$\frac{1}{k}\Big(\log Y_{i,j,t+k}) - \log Y_{i,j,t}\Big) = \alpha_0 + \beta(k)\eta_{i,t} + \delta Z_{t,j} + \varepsilon_{i,j,t}, \qquad k = 10, 20 \text{ years.} \tag{16}$$

The dependent variable $Y_{i,j,t}$ now represent the share of total non-farm employment for occupation i in industry j. The vector of controls $Z_{t,j}$ now contains time, or industry–time dummies depending on the specification, as well as lagged values of the dependent variable at the industry–occupation cell. The inclusion of industry–time allows us to isolate our findings from industry-specific trends in the sample.

Examining Panel B of Table 1, we see that the estimated slope coefficient on our innovation measure is consistently negative and economically and statistically significant. Overall, a one-standard deviation increase in $\eta_{i,t}$ is associated with a 0.60% to 0.89% decline in employment over the next 20-year horizon. The fact that this negative relation is essentially unaffected by the inclusion of industry-time fixed effects illustrates that our findings are not merely driven by the decline of certain industries which happen to employ workers with high technology exposure. Rather, much of the negative employment effects exist within, rather than between, industries. That said, the fact that the coefficients do attenuate slightly when including industry fixed effects indicates that high $\eta_{i,t}$ occupations tend to be employed in industries that have experienced overall employment declines.

Technology exposure, employment, and wages (1980–present)

We next turn our attention to the post-1980 period. We estimate the following specification

$$\frac{1}{k}\Big(\log Y_{i,t+k}) - \log Y_{i,t}\Big) = \alpha + \beta(k)\eta_{i,t} + \delta Z_{i,t} + \varepsilon_{i,t}, \qquad k = 5 \ldots 20 \text{ years.} \tag{17}$$

Here, $Y_{i,t}$ represents wage earnings or employment for a given occupation$i$ in calendar year t. The vector of controls $Z_{i,t}$ includes three lagged one-year growth rates of the dependent variable and time fixed effects. As before, $\eta_{i,t}$ is normalized to a unit standard deviation.

Figure 8 plots the estimated coefficients β along with 90% confidence intervals. We note that the responses for both wages and employment are strongly negative for all horizons. The point estimates are both economically and statistically significant and are comparable across horizons, suggesting these are permanent effects. Focusing on employment changes, a one-standard deviation increase in our technology exposure is associated with approximately a 1.1% annualized decline in occupation employment over the next five to twenty years. Similarly, our innovation measure also predicts a significant decline in wage earnings: a one-standard deviation increase in $\eta_{i,t}$ is followed by a decline in average wage earnings of approximately 0.2% to 0.3% per year over the same period.

Comparing the magnitude of employment declines in the MORG sample to the long-run (Census) results in Table 1 above, we note that the coefficient magnitudes are largely comparable. This is noteworthy in lieu of the fact the nature of breakthrough innovations in the post-1980 period is somewhat distinct than in the pre-1980 sample. As we saw in Figure 4, recent innovations are significantly more related to occupations performing routine cognitive tasks.

Recent work has argued that recessions are periods of technological transformation and thus accelerated automation of routine jobs (Jaimovich and Siu, 2018; Kopytov, Roussanov, and Taschereau-Dumouchel, 2018; Zhang, 2019). Consistent with this view, we next provide some direct evidence using our measure of technology exposure. In Figure 9 we see that occupations which were in the top quintile of $\eta_{i,t}$ in 1985 experienced stark declines in employment around the 1991, 2001, and 2007-2008 recessions, with a flatter but slightly declining profile in between recessions. Meanwhile, assigning occupations into the top quintile of routine-task intensity in 1985, we do see a persistent decline over the time period, but a much less pronounced pattern around recessions. This pattern is consistent with models of innovation-related job displacement where the opportunity to replace labor with an automation technology is a real option for firms. For example, Zhang (2019) shows that in a production based asset pricing model where firms choose to invest in labor-saving technologies, they choose to exercise this option when expected cash flows are temporarily low. Therefore the pattern exhibited in Figure 9 is consistent with employers replacing high $\eta_{i,t}$ workers with capital when the exercise value for doing so is high.

3.2 Individual-level evidence from panel administrative earnings data

Next, we turn our attention to individual worker outcomes. Doing so allows us to not only more directly link innovation to specific worker outcomes, but it also allows us to examine how this relation varies with observable worker characteristics. In particular, the detailed nature of administrative data allows us to examine how the relation between innovation and worker outcomes varies with proxies for worker skill, such as education or past earnings.

Data

We use a random sample of individual workers tracked by the Current Population Survey (CPS) and their associated Detailed Earnings Records from the Census—which contains their W2 tax income. The CPS includes information on occupation as well as demographic information such as age and gender. We limit the sample to individuals who are older than 25 and younger than 55 years old and to periods where the CPS interview date is less than 5 years old so that the occupation information is relatively recent.

We construct a measure of forward looking wage earnings growth following Autor, Dorn, Hanson, and Song (2014); Guvenen, Ozkan, and Song (2014); Kogan et al. (2020)

$$g_{i,t:t+h} \equiv w^i_{t+1,t+h} - w^i_{t-2,t}. \tag{18}$$

where $w^i_{t+1,t+h}$ refers to average *age-adjusted earnings* over the period, defined as

$$w^i_{t,t+h} \equiv \log\left(\frac{\sum_{j=0}^{h} \text{W-2 earnings}_{i,t+j}}{\sum_{j=0}^{k} D(\text{age}_{i,t+j})}\right). \tag{19}$$

refers to worker earnings net of life cycle effects. We focus on horizons of $h = 3, 5$, and 10 years. Appendix A.5 provides more details on the construction of the data and the patch to patent information.

Given that the administrative data sample has a shorter time dimension and a larger cross-section than our other data, we make some modifications to our construction of our worker technology exposure measure η. First, we identify important patents based on their 5-year, as opposed to 10-year forward similarity. This allows us to extend the sample by five additional years, which helps with the short length of the sample. Further, to fully take advantage of the larger cross-section, we allow our baseline innovation exposure measure η to also vary by industry, by restricting attention to patents issued to firms in the same 4-digit NAICS industry as the worker. Letting j index patents as before; $\Gamma_{k,t}$ denote the set of patents issued in industry k in year t; $o(i)$ the occupation of individual i; and $k(i,t)$ the industry of individual i in year t, we redefine our time series index of exposure of worker i to technology at time t as

$$\eta_{i,t} = \frac{1}{\kappa_t} \sum_{j \in \Gamma_{k(i,t),t}} \tilde{\rho}_{o(i),j} \times \mathbf{1}(\tilde{q}_{j,t} \geq \tilde{q}_{p90}). \tag{20}$$

In brief, our technology exposure metric (20) is largely the same as before, except that we now focus only on breakthrough patents in the industry in which worker i is currently employed—that is, $\eta_{i,t}$ defined in (20) now varies by occupation, industry, and year instead of just occupation and year as (12).

Given these restrictions imposed by the Census-CPS merged file and the nature of our innovation data, we are left with approximately 1.2 million person-year observations spanning the period from 1988 to 2016. In terms of demographics, approximately 58% of the sample is male and 46% of the observations correspond to workers with a college degree. Table 2 provides more details on the distribution of age and worker earnings: the median worker in the sample is 41 years old and earns approximately \$58k per year. The distribution of earnings is rather skewed however: the average is equal to \$75k while the 5th and 95th percentiles are equal to \$14k and \$172k, respectively. The last set of rows of Table 2 summarize the distribution of our (age-adjusted) cumulative earnings growth (18). At a horizon of $h = 5$ years, the median is equal to 0.011 while the mean is -0.063; given that (18) corresponds to a log difference, the large dispersion in earnings induces the mean growth rate to be negative due to Jensen's inequality. That said, the distribution is also highly negatively skewed: the 5th percentile is equal to -0.968 log points while the 95th percentile is equal to 0.574.

An illustrative example

Before examining the correlation between our technology measure and subsequent wage growth, it is informative to look at particular examples. We choose the rise of e-commerce—and more specifically the automatic fulfillment of retail purchase orders. Advances in information technology and telecommunications have obviated the need for manual processing of customer orders. Using our patent-based indicators we can identify the 1996 to 2002 period as a period of significant innovation related to the tasks performed by order-fulfillment clerks. Examples of such breakthrough innovations early on include U.S. Patent 5,696,906 for "Telecommunication user account management system and method"; Patent 5,592,560 for "Method and system for building a database and performing marketing based upon prior shopping history"; or Patent 5,628,004 for "System for managing database of communication of recipients". Appendix Table A.9 contains a longer list—the top 10 most related breakthrough patents to order fulfillment clerks issued in the 1997 to 2000 period.

To study the impact of these breakthroughs on worker earnings, Figure 10 contrasts labor market outcomes for order fulfillment clerks versus a set of workers in two related occupations unlikely to be affected by these innovations: personnel and library clerks.[12] The top panel of Figure 10 plots the average real wage (in 2015 US dollars) differential across the two sets of workers—normalized to be zero in 1997. The bottom panel plots the difference in average technology exposure from (20) for workers employed as order clerks across industries relative to workers employed as personnel or library clerks.

Examining the top panel, we note that relative wage trends for the two occupations were fairly flat prior to 1997. However, since then they begin to systematically diverge. The bottom panel shows that this divergence coincided with significant breakthrough innovations that were related more to order than library clerks. Beginning in 1996, there is a sustained increase in (relative)

[12]The DOT indicates that an order clerk 'Processes orders for material or merchandise received by mail, telephone, or personally from customer or company employee, manually or using computer or calculating machine...informs customer of any information needed...using mail or telephone. Writes or types order form, or enters data into computer, to determine total cost for customer. Records or files copy of orders."

innovation that persists for several years. By 2007, order clerks' average annual wages had declined by nearly $30,000 relative to personnel/library clerks.

Our preferred interpretation of the patterns in Figure 10 is that improvements in the automatic processing of orders displaced workers whose primary task was to fulfill orders. Naturally, we cannot exclude the possibility that these patterns reflect industry- or occupation-specific trends. Fortunately, our empirical design outlined below allows us to control for both industry- as well as occupation-specific year fixed effects.

Innovation and worker earnings growth

We estimate the following specification,

$$g_{i,t:t+h} = \alpha + \beta(h)\,\eta_{i,t} + \delta Z_{i,t} + \varepsilon_{i,t}. \tag{21}$$

Here, i now refers to a particular worker; as such, the left hand side represents the growth in her average earnings over the next h years compares to the last three. The main variable of interest $\eta_{i,t}$ now refers to the workers' technology exposure, specifically, the level of breakthrough innovations that are closely related to her occupation that are filed by firms in the same industry (based on NAICS 4-digit codes) as the worker. We examine earnings responses over the next $h = 3, 5$ and 10 years. The vector Z includes a rich set of controls that aim to soak up ex-ante worker heterogeneity. Specifically, we include various combinations of year, occupation and industry fixed effects—our most conservative specification interacts the latter two with calendar year to account for occupation- or industry-specific time trends. In addition, we include flexible non-parametric controls for worker age and past worker earnings as well as recent earnings growth rates.[13] Standard errors are clustered at the industry (NAICS 4-digit) level.

Table 3 summarizes our findings for the average worker in our sample. Overall, we find that workers' technology exposure is negatively related to their subsequent earnings growth. Panel A reports the estimated slope coefficients $\beta(h)$ for horizons of $h = 3, 5, 10$ years; different columns correspond to different fixed effect combinations. The magnitudes are both economically and statistically significant. Focusing on the 5-year horizon and the most conservative specification that includes both industry-year and occupation-year fixed effects, we see that a one standard deviation increase in innovation is associated with a 0.013 log point decline in average worker earnings over

[13] We construct controls for worker age and lagged earnings $w^i_{t-4,t}$ by linearly interpolating between 3rd degree Chebyshev polynomials in workers' lagged income quantiles within an industry-age bin at 10-year age intervals. In addition, to soak up some potential variation related to potential mean-reversion in earnings (which could be the case following large transitory shocks), we also include 3rd degree Chebyshev polynomials in workers' lagged income growth rate percentiles, and we allow these coefficients to differ by gender as well as past income levels based on five gender-specific bins formed based upon a worker's rank relative to her peers in the same industry and occupation.

the next five years. These magnitudes increase with the horizon h, ranging from 0.011 to a 0.014 log point decline in average earnings at horizons of three and ten years, respectively.[14]

In brief, Panel A shows that the average effect of technology exposure on worker earnings is negative. This negative average effect, however, likely masks considerable heterogeneity in ex-post worker outcomes. To that end, in Panel B we next examine whether technology exposure is associated with increases in the second-moment of earnings growth: we estimate a modified version of (21), in which now the dependent variable is the *absolute value* of earnings growth. We see that technology exposure is associated with an increase in the second moment: the absolute value of earnings growth (21) increases by approximately 0.5 percentage points in response to a one-standard deviation increase in $\eta_{i,t}$, which corresponds to approximately a 3% increase relative to the sample median value of the dependent variable (0.157).

Similarly, Panel C examines the increase in the skewness of the earnings distribution, or equivalently the extent to which the negative average effects documented in Panel A are concentrated in a subset of workers. Specifically, we construct indicators for whether a given worker's income growth over a given horizon falls in the bottom 10th percentile of all our observations within a given year. Focusing on the same specification, we see that an increase in a worker's technology exposure is associated an economically significant increase in the likelihood of large earnings losses for affected workers: a one-standard deviation increase in $\eta_{i,t}$ is associated with a 0.4 percentage point increase in the likelihood that a worker's subsequent earnings growth is in the bottom 10th percentile.

Our findings in this section reinforce our findings in Section (3.1) that technology is associated with occupation-level declines in employment and wages. By tracking the earnings growth of individual workers, we can ensure that our findings on worker earnings are not driven by selection across occupations. In addition, individual-level data allows us to paint a more complete picture of how earnings losses are distributed across workers, even at a particular industry–occupation cell at a point in time. Building on this idea, the next section examines how these findings vary with worker characteristics.

Innovation, worker earnings and ex-ante heterogeneity

We next allow the effects to vary by observable worker characteristics. In terms of the heterogeneous effects of technological progress across workers, existing work has emphasized that (a) much of technological change is skill biased (see, e.g. Goldin and Katz, 2008, for a textbook reference); (b) an important component of human capital is likely specific to a technological vintage (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). Accordingly, we focus on education

[14] We find no meaningful differences across genders: over the next five years, on average men experience a 0.013 log point decline and women experience a 0.011 log point decline.

and prior income as common proxies for worker skill; allowing the impact of technology to vary by worker age helps tease out the effect of vintage-specific human capital. In what follows, we re-estimate equation (21) and now allow the slope coefficient $\beta(h)$ to vary across worker sub-groups. For brevity we focus on the most conservative specification that includes both industry-year and occupation-year fixed effects. Appendix Tables A.11 to A.12 illustrate that results are similar across alternate specifications.

Table 4 examines how the effect of technology on worker earnings vary with education, specifically on whether the worker has a college degree. Focusing on mean growth rates across horizons—columns (1) to (3)—we see that an increase in technology exposure $\eta_{i,t}$ is associated with an economically and statistically significant decline in average earnings growth for both college and non-college educated workers. If we interpret education as a proxy for skill, the fact that earnings decline for both groups is somewhat at odds with the canonical model of skill-based technical change. That said, we do find that non-college workers experience a somewhat larger decline in average earnings growth than workers with a college degree, with the difference being marginally statistically significant (p-values range from 0.043 to 0.11 across horizons). Columns (4) and (5) show that both groups experience a similar increase in their second moment of earnings growth, while non-college educated workers experience a somewhat larger increase in the probability of large earnings declines than college educated workers (0.48 vs 0.35 percentage points).

Table 5 examines how the response of worker earnings growth to $\eta_{i,t}$ varies by prior income. In terms of point estimates, we find a U-shaped pattern: a given increase in a worker's technology exposure $\eta_{i,t}$ has the largest impact on the earnings growth of not only the least- but also the highest-paid workers (relative to their peers in the same occupation and industry). These estimates are noisy: we cannot reject the null that the response of earnings growth of workers at the bottom quartile is equal to the responses of workers in the 25th to 95th percentile. However, the response of the most-highly paid workers is both statistically as well as economically distinct from the workers in the middle group: a one-standard deviation increase in technology exposure is associated with a 0.025 log point decline in their average earnings growth—approximately twice as large as the effect on the average worker.

At face value, these facts seem at odds with the canonical model of skill-biased technical change: if technology is complementary to the labor input of skilled workers, to the extent that prior income is related to worker skill, one would expect to see that top workers should experience an increase in their average earnings (or at the very least a smaller decline). By contrast, the opposite pattern obtains. Comparing across columns (1) to (3) we see that these patterns are quantitatively similar across horizons, suggesting that these effects are highly persistent.

One potential reconciliation of the patterns in Table 5 with the standard view of technology-skill complementarity is allowing for vintage-specific human capital. That is, skill is not an immutable characteristic of the worker; it is the result of experience and learning by operating a particular technology. When new technologies are introduced, some of that accumulated knowledge becomes obsolete: skilled workers in the old technology need not remain skilled in the new. If that is the case, we would expect skilled workers to face greater earnings risk in response to increased rate of technological innovation due to the possibility of skill displacement. Consistent with this view columns (4) and (5) of Table 5 show that top earners face significantly greater labor income risk than the average worker in response to an increase in their technology exposure. Focusing on the last column, a one-standard deviation increase in $\eta_{i,t}$ is associated with a 1.26 percentage point increase that these workers experience a large earnings decline (earnings growth in the bottom 10-the percentile) which is approximately three times higher than the average worker.

Table 6 provides additional support to the idea of vintage-specific human capital by examining how these effects vary with worker age. Older workers are both more likely to have accumulated skills in existing technology and also less likely to be able to become familiar with new production methods. Accordingly, we see that older workers (those in the 45 to 55 range) experience significantly greater declines in earnings growth (0.02 to 0.025 log points across horizons) relative to workers aged 35–45 (0.9 to 1.4 log point decline) or 25–35 (0.4 to 0.9 log point decline). Columns (4) and (5) similarly show that earnings risk in response to an increase in technology exposure is increasing in age: a one-standard deviation increase in $\eta_{i,t}$ is associated with a 0.9 percentage point increase in the likelihood of a large earnings decline for workers aged 45–55, compared to a 0.2 percentage point increase for younger workers.

Discussion

In brief, we find that a given improvement in technology leads to lower earnings growth across all workers. These patterns, together with the positive correlation to industry productivity and the decline in the labor share documented in Section 1 is consistent with the view that most of the breakthrough innovations in the sample are labor-saving, that is, they partly replace tasks performed by workers. The standard view is that increases in automation are more likely to affect low-skill workers, since low-productivity workers are likely to be replaced first. Some of our findings are consistent with this view: we find some evidence that workers without a college degree and the lowest-paid workers experience larger than average earnings declines in response to technological innovation (though the latter difference is not statistically significant).

However, some of our findings are harder to reconcile with a view that skill is an immutable characteristic of the worker. Workers that are more highly paid relative to their peers in the same occupation and industry experience on average significantly larger earnings declines than the average worker; further, our findings suggest that much of these average decrease reflects an increase in the probability of large earnings losses as opposed to a small consistent decline in earnings. Put differently, the distribution of earnings losses is heterogenous: a subset of skilled workers experience large earnings declines rather than the entire group experiencing small declines.

This pattern, together with the increased earnings response of older workers suggests a role of vintage-specific human capital, or equivalently for technology making certain worker skills obsolete. The model in the next Section 4 formalizes and quantifies this idea more fully. That said, one caveat in interpreting these patterns is that the period covered by the Census-CPS administrative data coincides with the rise of very specific technologies, namely ICT. Thus, we should be careful when extrapolate these findings to other periods of rapid technological progress.

3.3 Additional Results and Robustness Checks

Here we discuss a number of additional results that frame our work to the existing literature and explore the extent to which our measure underestimates the degree of labor-displacive innovations.

Comparison to Existing Measures of Exposure to Technical Change

Our work is not the first to construct occupation-level measures of exposure to technological change (Autor and Dorn, 2013; Webb, 2019). A key advantage of our measure relative to existing work is that it also incorporates time-series variation. That said, it is instructive to explore the extent to which it contains additional information regarding cross-sectional differences in technology exposures. Here, we explore this question, and compare the performance of our technology measure in predicting wage and employment declines relative to the routine-task intensity and the measure of occupation-level offshorability from Autor and Dorn (2013), and the Webb (2019) measures of exposure to robotics or software.

To compare our different approaches, we estimate a long-difference cross-sectional specification similar to Webb (2019) as follows

$$\frac{1}{k}\Big(\log Y_{i,t+T}) - \log Y_{i,t}\Big) = \alpha + \alpha_j + \beta \eta_{i,1980} + \delta X_i + \epsilon_{i,j} \tag{22}$$

Here i indexes occupations and j indexes industries. In estimating (22), we combine information on wages and employment in the 1980 Census and the 2012 ACS. In particular, we use the 1980 Census and 2012 ACS data from Deming (2017), which are reported at the occupation by industry by education level, and aggregate the data to industry by occupation. Thus, the dependent variable denotes either the log change in employment or the change in log wages over the 1980-2012 time period. We include industry fixed effects α_j to account for industry specific shocks that may be correlated with occupational outcomes. Controls X_i include occupation employment share in 1980, occupation log wage in 1980, three indicators for the occupations education level in 1980, the routine-task intensity and the measure of occupation-level offshorability from Autor and Dorn (2013), and the Webb (2019) measures of exposure to robotics or software patents, depending on the specification. We weight observations by the employment share in 1980 and cluster standard errors by industry.

Tables 9 reports our findings for employment (Panel A) and wages (Panel B). Examining the first row of each panel, we see that the point estimates (and statistical significance) of β are essentially unaffected by including the Autor and Dorn (2013) or Webb (2019) measures. We conclude that cross-sectional differences in $\eta_{i,t}$ contain independent information relative to these alternate cross-sectional metrics.

Alternative approaches to measuring labor displacement

Our technology exposure measure is constructed based on the similarity between patents and tasks performed by a specific occupation. Ex-ante, it is not entirely obvious whether a high level of similarity is likely to capture complementarity or substitution between the technology and the tasks performed by labor. Even though we are finding a consistently negative relation between our technology exposure measure and subsequent labor market outcomes, it is possible these effects are muted because our measure mixing labor-saving and labor-enhancing innovations.

To explore this possibility, we next compare the performance of our measure to a purely statistical predictor that is calibrated to predict employment declines in-sample. To do so, we leverage recent advances in topic modeling to construct a composite predictor from patent text whose purpose is to maximize the in-sample predictability of employment declines. This measure is akin to a principal component; it has no straightforward economic interpretation, but it rather provides a statistical upper bound on how large the labor-displacive effects could be on a factor that is constructed from the text of breakthrough patents.[15] The correlation between our baseline measure $\eta_{i,t}$ and the statistical predictor constructed to represent exposure to labor-saving technologies is approximately 73 percent.

We then compare the performance of our baseline measure based on patent-task similarity to the in-sample performance of this statistical factor in the wage and employment regressions in Figure

[15] We build on the approach proposed by Cong, Liang, and Zhang (2019), which is well-suited to prediction exercises using large-scale textual data. In brief, this approach can be summarized as follows. We first extract the 500 most important common factors (topics) from the text of breakthrough patents using the approach of Cong et al. (2019) and the vector representations of word embedings discussed in Section 2. We then use these 500 textual factors to form a single predictor that is optimized to predict occupation declines in-sample. To do so, we examine the univariate performance of each factor in predicting employment declines, and then form a linear combination (the first principal component) of the predictors that are statistically significant negative predictors at the 5%. We also construct a labor-enhancing factor using the converse exercise. Appendix **??** describes the procedure in detail.

8—which can be viewed as an upper bound in the ability of patent text to predict employment and wages. We find that the performance of our baseline measure is close to this upper bound: the annualized employment and wage declines predicted by this statistical displacement factor are 1.25% and 0.25%, respectively—compared to 1.12% and 0.20% for our technology measure based on patent-task similarity. Appendix **??** provides more details. We conclude that our technology measure based on patent-task similarity primarily captures labor-saving innovations.

Robustness

We perform several checks to explore the extent to which our findings are specific to a particular period. Appendix Table A.5 shows that our findings on the long-run negative relation between technology exposure and employment at the occupation level are robust across sub-samples. That said, Appendix Table A.6 shows that this negative relation is primarily driven by innovation waves: we separate the sample into two sub-samples, where we define as innovation waves the 20 year periods beginning in years 1880, 1910, 1920, and 1980, 1990—with the remaining years representing non-innovation wave periods (for a discussion of the 1920s, see e.g. Field, 2003).

Further, we verify the robustness of our findings to alternative specifications. Appendix Tables A.11 through A.13 show that our results on heterogenous responses by education, age, and prior income are largely robust to different combinations of fixed effects and horizons over which we measure worker earnings.

That said, one particular caveat in interpreting the heterogenous patterns we uncover using the Census-CPS administrative data is that they are estimated during a very specific time period which coincides with the rise of very specific technologies (i.e. ICT). The extent to which similar patterns obtain more broadly during earlier periods is an open question. As a step towards this direction, we re-estimate our long-run employment results and allow the coefficient to vary by worker age (the only variable consistently reported since the 1850s in the population sensus). In particular, we re-estimate 15 but now add a second panel dimension, worker age. That is, the unit of analysis is not worker occupation-age groups and when we compute employment growth rates we track a particular cohort. For instance, to compute the 20-year growth rate in employment in 1900 for workers aged 20–29 in occupation i, we compare it to the employment of 40–49 year old workers in occupation i in 1920.

Table A.14 reports our findings. Panel A focuses on the full sample. We see that the employment growth of older workers in a specific occupation is significantly more exposed than the employment growth of younger workers (difference has a p-value of 0.015). In terms of magnitudes, a one-standard deviation increase in technology exposure is followed by a 1.1% annual decline in employment for older workers over the next twenty years, compared to a 0.7% annual decline for younger workers.

However, Panel B shows that this pattern is largely driven by the latter part of the sample. Specifically, there are no differences in employment outcomes across age groups during the 1850–1920 period. In the 1930 to 1960 period, there is some evidence that older workers are significantly more exposed, but the results are too noisy to infer meaningful differences (p-value of 0.13). By contrast, focusing on the post-1970 period, the difference is between younger and older workers becomes larger and statistically significant. We conclude that we cannot rule out the possibility that our worker heterogeneity results are somewhat specific to the ICT revolution: it is entirely possible that older workers were significantly more displaced than younger workers by ICT.

4 Model

Here we provide a model that features skill-biased technological change and allows for skill displacement. The model contains a continuum of workers who supply high- and low-skill labor inputs. Consistent with the literature, the output of the high-skill labor input is more complementary to technology than the output of the low-skill input. As a result, improvements in technology lead to an increase in the wages of high-skill workers relative to the wages of low-skill workers. This is a feature of the standard model.

We extend the model to allow for skill displacement. In particular, improvements in technology may render some of workers' skills obsolete—a specific worker may lose a part of her skill when the technology frontier improves. As a result, even though wages of skilled workers rise in response to technology, an incumbent skilled worker may experience a decline in wages.

4.1 Setup

We model the output of a given industry. Output is produced by three factors of production: low-skilled labor L, high-skilled labor H, and intangible capital (technology) ξ. For simplicity, we will abstract from labor growth and model output per capita Y as

$$Y_t = \left[\mu\Big(H_t\Big)^{\sigma} + (1-\mu)\left(\lambda\Big(\xi_t\Big)^{\rho} + (1-\lambda)\Big(L_t\Big)^{\rho}\right)^{\sigma/\rho}\right]^{1/\sigma} \tag{23}$$

Here, ρ denotes the elasticity of substitution between technology and unskilled labor and σ denotes the elasticity of substitution between skilled labor and the composite output of technology and unskilled labor.[16] Since the total mass of workers is normalized to one, equation (23) also refers to

[16]In setting up (23), we have included technology and unskilled labor in the inner nest, and skilled labor in the outer nest. This is in contrast to the formulation in Krusell et al. (2000); Eisfeldt, Falato, and Xiaolan (2021), but note that the comparison with these two papers is imperfect, since ξ denotes intangible capital (technology) rather than physical equipment (machines).

labor productivity (output per worker).

The factor ξ is the stock of intangible capital/knowledge embodying the technology used for producing output Y, similar in spirit to Acemoglu and Restrepo (2018). When we map our empirical analysis to the model, we will interpret our technology exposure metric $\eta_{i,t}$ as a shock to ξ, which however affects only a subset of workers involved in the production of Y—the model equivalent to 'occupations', described below. Keeping with the literature, we expect technology to be more complementary to skilled labor relative to unskilled labor, so we will impose the condition that, in relative terms, shifts in technology are more complementary to skilled than unskilled labor, or equivalently that

$$\sigma < \rho < 1. \tag{24}$$

Put differently, technology ξ is a better substitute for unskilled rather than skilled labor.

Technology ξ evolves exogenously according to

$$d\xi_t = -g\,\xi_t\,dt + \kappa\,d\,N_t. \tag{25}$$

Technology improves according to the process N_t whose increments are Poisson with arrival rate $\omega\,dt$. Recall that we have set up output in per capita terms. As such, the negative drift term in equation (25) reflects the fact that ξ also a per-capita quantity and population grows at rate g. Given (25), the level of ξ is stationary with a long-run mean equal to $\kappa\,\omega/g$.

Workers are heterogenous along two dimensions. In particular, there is a unit mass of workers differentiated by their type $\theta \in [0, 1]$, which determines their endowment of high- and low-skill labor inputs; workers also vary in their ability to acquire new skills $s = \{l, h\}$. Specifically, each worker can provide θ units of skilled labor H and $1 - \theta$ units of unskilled labor L. As a result, the total supply of skilled labor as a share of population is equal to

$$H_t = \int_0^1 \theta\, p_t(\theta)\, d\theta, \tag{26}$$

where $p_t(\theta)$ is the measure of workers of skill level θ at time t. Since we normalize the total supply of labor to one,

$$L_t = 1 - H_t. \tag{27}$$

In addition, workers vary in their ability to acquire new skills—that is, increase their skill level θ. The share of workers who cannot acquire new skills s_l can produce only in the low-skill task so $\theta = 0$. The remaining share of workers $s_h = 1 - s_l$, have skill $\theta \in (\underline{\theta}, 1)$ that evolves over time due to learning by doing and technological displacement according to

$$d\theta_{i,t} = m\,\theta_{i,t}\,dM_{i,t} - h\,\theta_{i,t}\,d_{i,t}\,d\,N_t, \tag{28}$$

Here, $dM_{i,t}$ is a Poisson jump with arrival rate $\phi\,dt$ that reflects the stochastic acquisition of new expertise. Since we limit $\theta \in (\underline{\theta}, 1)$ for these workers, we impose reflecting boundaries at $\underline{\theta}$ and 1.

Importantly, the last term in equation (28) captures the displacive effect of the arrival of new technologies ($d\,N_t = 1$). There is a stochastic element in how technology improvements affect workers: this uncertainty is captured by $d_{i,t}$, which is a random variable with support on the unit interval and is independent of θ_{it}. For now, we assume that $d_{i,t}$ i.i.d. distributed across agents and follows a binomial distribution $d \in \{0,1\}$ with $Prob(d = 1) = \alpha$. More generally, we could allow the distribution of d_i to vary with certain worker characteristics such as age or education. Affected workers experience a proportional loss in their human capital (skill) by a factor h. Last, workers of each type die at Poisson rate $\delta\,dt$ and are replaced by newborn skilled workers with either zero skill ($\theta = 0$) or the minimum level of skill ($\theta = \underline{\theta}$) for skilled workers with probabilities s_l and $1 - s_l$ respectively.[17]

Given our assumptions (26)–(28), the aggregate supply of skilled labor H_t increases with learning, decreases as skilled older workers are replaced with unskilled young workers, and decreases temporarily following periods of rapid technological progress. The latter effect captures the idea that technological improvements may be associated with lower output in the short run as agents in their economy need to upgrade their skills to fully take advantage of new innovations—similar in spirit to Brynjolfsson, Rock, and Syverson (2018).

The current wage of an individual worker with skill level $\theta_{i,t}$ is equal to

$$w_{i,t} = W_{L,t} + \theta_{i,t}\left(W_{H,t} - W_{L,t}\right). \tag{29}$$

In equilibrium $W_{H,t}$ and $W_{L,t}$ are equal to the marginal product of skilled and unskilled labor, respectively

$$W_{H,t} = \frac{\partial Y_t}{\partial H_t}, \quad \text{and} \quad W_{L,t} = \frac{\partial Y_t}{\partial L_t}. \tag{30}$$

In sum, we provide a model in which the skill premium increases with the level of technology, yet the wage earnings of individual skilled workers can fall as they potentially are displaced. We discuss the model calibration next.

[17] Our formulation for θ is related to Jones and Kim (2018) in that the skill of an individual worker grows on average over time but occasionally resets to a lower level.

4.2 Model Calibration

Here we discuss how we fit the model to the data.

Methodology

The model has a total of 14 parameters. We choose these parameters via a mixture of calibration and indirect inference. Specifically, we choose $s_l = 0.375$ so that workers with only low-skill labor inputs constitute the lowest income bin (25% of the sample), and half of the second-lowest. Since m and ϕ are not separately identified, we set the learning rate $m = 0.03$; when choosing the grid for θ, we assume that skilled workers human capital $\theta \in (0.03, 1)$. Last, we set the worker exit rate, at $\delta = 2.5\%$ which corresponds to a 40 year average working life.

To estimate the remaining 10 parameters $\Theta = \{\mu, \lambda, \rho, \sigma, \phi, \alpha, \kappa, \omega, h, g\}$, we target the mean level of the skill premium, the response of labor productivity and the labor share to changes in technology estimated in Section 1, and the response of worker earnings growth and likelihood of large wage declines conditional on levels of prior income, estimated in Section 3.2. Since the model has no mechanism for delayed responses, whereas in the data the diffusion of technology likely takes some time, we match the model responses on impact to the empirical responses over five years. Table 8 summarizes the 14 statistics that we target.

To obtain some intuition for how the model parameters are identified, we next discuss how these quantities help identify model parameters. In the model, the skill premium defined as the ratio of wages for the high-skill versus the low-skill labor input equals

$$\frac{W_{H,t}}{W_{L,t}} = \frac{\mu}{(1-\mu)(1-\lambda)} \left(\frac{H_t}{L_t}\right)^{\sigma-1} \left(\lambda\left(\frac{\xi_t}{L_t}\right)^{\rho} + (1-\lambda)\right)^{\frac{\rho-\sigma}{\rho}}. \tag{31}$$

Examining (31), we see that as long as technology is more complementary to high-skill than low-skill labor inputs ($\rho > \sigma$) and we hold fixed the supply of skilled and unskilled labor H and L then the skill premium is increasing with the level of technology ξ. Thus, just like the standard model, our model generates an increase in the skill premium over the long run during periods when the rate of technology rises faster than average. However, in our model the supply of high- and low-skill inputs H and L varies in the short run, due to skill displacement (28). This process leads to drop in H/L and thus a further increase in the skill premium in the short-run.

When mapping the model to the data, we define the skill premium as the mean ratio of earnings of workers in the 75th vs the 25th percentile. This ratio combines information on the ratio W_H/W_L and the ergodic distribution of θ. In terms of identifying model parameters, the mean level of the skill premium thus helps identify the factor share parameters μ and λ and the elasticities ρ and σ.

Further, it affects the parameters driving the ergodic distribution of θ, namely ω, ϕ, h and α.

The labor share of output in the model can be written as

$$\frac{W_{H,t}\,L_t + W_{L,t}\,H_t}{Y_t} = \frac{(1-\lambda)\,(1-\mu)\left(\lambda\,\left(\frac{\xi_t}{L_t}\right)^{\rho} + 1 - \lambda\right)^{\frac{\sigma}{\rho}-1} + \mu\,\left(\frac{H_t}{L_t}\right)^{\sigma}}{(1-\mu)\left(\lambda\left(\frac{\xi_t}{L_t}\right)^{\rho} + 1 - \lambda\right)^{\frac{\sigma}{\rho}} + \mu\,\left(\frac{H_t}{L_t}\right)^{\sigma}} \tag{32}$$

The response of the labor share (32)— and output (23)—to increases in technology ξ is ambiguous in the model. These both depend on the extent to which different tasks contribute to output (μ and λ); technology-labor complementarity (ρ and σ) and the response of H and L to a technology shock (which depends on h and α). What helps with identification in our case is our finding that technology improvements are associated with declines in the labor share of output and an increase in output per worker (see Section 1). The fact that output/productivity and the labor share respond with opposite signs helps narrow down the set of admissible parameters quite significantly.

To identify the parameters involved in the dynamics of worker skill acquisition and displacement in (28), we also target the heterogeneity in earnings responses to changes in technology (see Section 3.2). Specifically, we target the mean earnings growth responses (column (2) in Table 5), and changes in the probability of large declines in wage earnings (column (5) in Table 5). To construct the analogue. Recall that, in the data, there is a U-shape relation in mean responses, with the highest-paid workers and lowest-paid workers experiencing the largest earnings declines in response to technology shocks. In terms of higher moments though, we find that the highest-paid workers are far more likely than other groups to experience large earnings declines. In the model, whether higher-paid workers are more exposed to technology is largely ambiguous.

To see this, we can derive the following decomposition for wage earnings growth in the model over any horizon h:

$$\frac{w_{i,t+h} - w_{i,t}}{w_{i,t}} = \underbrace{\frac{w_{l,t}}{w_{l,t} + \theta_{i,t} s_{p,t}}}_{\substack{\text{low skill}\\ \text{income share}}} \underbrace{\frac{\Delta_h w_{l,t+h}}{w_{l,t}}}_{\substack{\text{low skill}\\ \text{wage chg}}} + \underbrace{\frac{\theta_{i,t} s_{p,t}}{w_{l,t} + \theta_{i,t} s_{p,t}}}_{\substack{\text{high skill}\\ \text{income share}}} \left[\frac{s_{p,t+h}}{s_{p,t}} \cdot \underbrace{\frac{\Delta_h \theta_{i,t+h}}{\theta_{i,t}}}_{\substack{\text{skill}\\ \text{displacement}}} + \underbrace{\frac{\Delta_h s_{p,t+h}}{s_{p,t}}}_{\substack{\text{high skill}\\ \text{wage chg}}}\right] \tag{33}$$

where $s_{p,t} = W_{h,t} - W_{l,t}$.

As we see from the last term in brackets in (33), whether the highest-earning workers experience larger declines depends on whether the increase in the skill premium in is sufficient to offset the loss of worker skill $\theta_{i,t}$ due to skill displacement—see equation (28). For the high-income (i.e. θ) workers, the primary income risk in the model comes from having human capital displaced, while the lowest-income workers (those in the s_l group with $\theta = 0$) face income losses from changes in wages. Improvements in technology lead to an increase in the skill price of H and a drop in the skill price of L because of both differences in complementarity and skill displacement—since workers fall down the ladder following a shock, H is scarcer and L is more abundant. These effects depend on the size of human capital losses and increases, as well as the associated skill prices following displacement, including h, ϕ, ω, λ, μ, σ, and ρ.

Mapping the empirical regressions in Table 5 to the model entails two challenges: first, our technology measure varies at the industry and occupation level whereas the model refers to a single industry; second, our empirical specifications include occupation, industry, and time fixed effects so the main coefficients are also identified by comparing to workers in other occupations or industries. To narrow the gap between the model and the data, we construct the closest equivalent to a regression coefficient in the model as follows. We first calculate a set of wage responses that vary by income bins that match the empirical equivalents. Within each income bin, we compute wage growth for exposed ($d_{i,t} = 1$) and unexposed ($d_{i,t} = 0$) workers in the case of a technology shock occurring ($dN_t = 1$) or not ($dN_t = 0$). The equivalent of the regression coefficient in the model is the coefficient of wage growth on the interaction between a shock occurring and the worker being exposed, while separately controlling for exposure and shock dummies and everything interacted with income bins.[18]

We calibrate the remaining 10 model parameters by minimizing the distance between the output of the model $\hat{X}(\Theta)$ and the data X,

$$\hat{\Theta} = \arg\min_{\Theta} \left(X - \hat{X}(\Theta)\right)' W \left(X - \hat{X}(\Theta)\right). \tag{34}$$

Our choice of weighting matrix W emphasizes percent deviations of the model vs the empirical values and places relatively more weight in the aggregate moments.

Table 7 summarizes our parameter choices. Similar to Krusell et al. (2000); Eisfeldt et al. (2021), we find that technology is a good substitute for the low-skill labor input ($\rho = 0.74$) whereas the high-skill labor input is complementary to technology ($\sigma = -0.12$). Technology shocks are relatively frequent ($\omega = 1.56$) and sizeable ($\kappa = 0.32$). Importantly, however, the model features a modest degree of skill displacement: workers who fall down the ladder only lose $h = 6\%$ of their existing level of θ. That said, these losses are pervasive: the probability of skill loss conditional on a shock is $\alpha = 32\%$. Further, these skill losses are transient: workers are able to acquire skills (increase θ) at an average rate of $m\,\phi = 7.2\%$ per year.

[18]When constructing these regression coefficients in the model, we use the ergodic distribution of wage growth, so we take into account the share of exposed workers α, the frequency of technology shocks ω and the likelihood each worker falls in a given income bin.

Model Fit and Discussion of the Mechanism

Examining Table 8, we see that the model does a good job matching the target statistics, including the labor share and responses of aggregate quantities to technology shocks. Specifically, the model is able to capture the fact that output and labor productivity rise following a technology shock whereas the labor share falls. In addition, the model is able to largely replicate both the marginal effect of a shock on exposed workers as well as the U-shape pattern of coefficients by income rank.

Figure 11 plots the impulse responses generated by the model in response to a one-standard deviation shock to the level of technology ξ (panel A). Panel B shows that this improvement in technology leads to a 2.5% rise in output/productivity on impact. By contrast, Panel C shows that the labor share declines by approximately 1.5%. This decline in the labor share is driven by a combination of two factors. First, as we see in Panel D, the quantity of the high-skill labor input declines by approximately 2.5% as workers' skills are displaced. This fall is temporary, as H gradually increases to skill acquisition. Since the wages for the high-skill task exceed the wages of the low-skill task, the total wage bill in the economy falls. Second, Panel E shows that improvements in technology are associated with decline in the price of the low-skill labor input (W_L) which further depresses the labor share; by contrast, even though the price of the high-skill labor input rises in Panel F, the rise is not sufficient to cause the labor share to rise because H falls. These movements in skill prices are driven by a combination of two forces: first, the high-skill input is complementary to ξ whereas the low-skill input is a substitute; second, skill prices change in response to the reduction in the effective supply of H due to skill displacement.

Figure 12 summarizes the distributional impact of technology shocks in the cross-section of workers. Panel A focuses on differences in growth rates in response to a technology shock relative to the no-shock counterfactual. The blue bars correspond to unexposed workers ($i.e. d_i = 0$). For these workers, the only effect in play is changes in skill prices. Low-income workers supply only the low-skill labor input L. Since the price W_L of the low-skill input falls, these workers experience a decline in wages. By contrast, the high-income workers supply mostly the high-skill input H; since the price of the high-skill input W_H rises, these workers experience an increase in wages. Absent skill displacement would be the only impact on wage growth in the model—and would be similar to the models in Krusell et al. (2000) and Eisfeldt et al. (2021). Yet, such a model would be unable to produce the empirical patterns in Table 5.

The orange bars in Panel A of Figure 12 correspond to the wage growth of exposed (i.e. $d_i = 1$) workers following a shock relative to the no-shock counterfactual. These workers experience the same change in skill prices as the unexposed workers, but they are also subject to skill displacement (loss of human capital θ). As a result, the wage growth of the high-income exposed workers is markedly different than the wage growth of the unexposed high-income workers: despite the fact that skill prices W_H rise, these workers experience a fall in wages due to loss of human capital θ. Further, just like the data, their wages fall significantly more than the low-income workers, implying that this loss in skill is significant.

Panel B Figure 12 plots the equivalent of the regression coefficient in the model, that is, the OLS coefficient of a regression of wage growth on a shock and exposure dummy, controlling for income. Since these slope coefficients are estimated using the ergodic distribution of wages at the model steady state, which factor in the relative size of the different worker groups and the frequency of technology shocks they cannot be expressed as simple functions of the coefficients in Panel A. However, they display a similar pattern as the orange bars: improvements in technology have an asymmetric effect on the wages of exposed workers. The workers most affected are the high-income workers—with some mild evidence of a U-shaped response.

In brief, Figure 12 summarizes the impact of technology of wages, which is a combination of shifts in skill prices and changes in the quantity of human capital. The combination of these effects generate the U-shape in earnings losses we see in column (2) of Table 5. The lowest-income workers have $\theta = 0$, and as a consequence have wages which fall dramatically relative to a non-shock period. Workers in the middle part of the income distribution experience some loss of human capital and suffer from the decline in the price of low-skill labor input W_L, but these losses are partly offset from the increases in the high-skill price W_L. Workers at highest income group has the farthest to fall: these workers who are exposed to technology experience the largest wage declines of anyone in the model due to skill displacement. By contrast, unexposed workers who stay at the top of the ladder following a technological innovation see large wage increases due to higher W_H—which results from scarcer H and the complementarity of H and ξ.

In addition to the impulse responses to a given shock holding the model parameters fixed, we can also study the impact of shifts in structural parameters. In Figure 13 we compute transition paths associated with a permanent increase in ω, the rate of technological innovation and therefore a permanent increase in the level of technology ξ. We calibrate the increase in ω so that the new steady-state level of ξ is one-standard-deviation higher than in our calibration. As we see in the blue line, this shift leads to a permanent increase in output and productivity, a decline in the labor share, an increase in the skill premium, and an increase in inequality. Further, the permanent increase in the rate of innovation also leads to a permanent decline in the quantity of the supply of the high-skill input due to high rate of displacement. As a result, the level of output eventually starts to drop back towards the initial steady state because H and ξ are complements. The orange line in the same plot shows how these responses vary if we also change the rate of skill acquisition ϕ to keep the steady-state level of H constant between the two steady states.

5 Conclusion

We develop a new method for identifying the arrival of labor-displacive innovations. Our time series indicators of worker technology exposure date are available since the mid 19th century and are available at a high level of granularity—industry and worker occupation. Examining the type of worker tasks most exposed to innovation, we find that while non-routine manual (physical) and routine-manual tasks have been highly exposed throughout the last 150+ years, the innovations of the information technology revolution in the post-1980 period saw an increased relationship with cognitive tasks.

More importantly, we find that our technology exposure measures are consistently negatively related to workers' future labor market outcomes, both at the group (occupation) but also at the individual level. Using a panel of administrative data on worker earnings, we show that the earnings of older and less educated workers are more responsive to our technology exposure measure, which is in line with the existing view of technology-skill complementarity. By contrast, our finding that the earnings of more highly paid workers (relative to their peers in the same industry and same occupation) respond more to our technology measure is somewhat at odds.

We reconcile these patterns with the standard view by allowing for skill displacement in the standard model of technology-skill complementarity. Our calibrated model is able to quantitatively replicate our main findings: improvements in technology are associated with increases in productivity but a decline in the labor share; lower earnings across workers of all income groups. In the model, the earnings of high-income workers respond more to technology improvements because these workers have further to fall: the loss of skills following technological progress is sufficient to offset any wage gains associated with higher skill prices.

Overall, we provide long-run evidence that the technological displacement of labor has been a persistent phenomenon over the past century and a half. Our findings illustrate the utility of our technology exposure indicators that can be used to study an array of questions in economics. That said, we should emphasize that our indicators are constructed largely from the perspective of incumbent workers and are primarily intended to capture technological substitution of existing tasks. A likely feature of technological progress that we are missing is that it facilitates the creation of new tasks and occupations. Building on our work, Autor, Salomons, and Seegmiller (2021) represents a promising step along that direction.

References

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

1043–1171.

Acemoglu, D., G. Gancia, and F. Zilibotti (2012). Competing engines of growth: Innovation and standardization. *Journal of Economic Theory 147*(2), 570–601.

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

Acemoglu, D. and P. Restrepo (2020). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy forthcoming.*

Acemoglu, D. and P. Restrepo (2021). Tasks, automation, and the rise in us wage inequality. Working Paper 28920, National Bureau of Economic Research.

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

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

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

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

Autor, D., A. Salomons, and B. Seegmiller (2021). New frontiers: The origins and content of new work, 1940–2018. Working paper, MIT.

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

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

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

Bianchi, F. (2016). Methods for measuring expectations and uncertainty in markov-switching models. *Journal of Econometrics 190*(1), 79–99.

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

Brynjolfsson, E., D. Rock, and C. Syverson (2018). The productivity j-curve: How intangibles complement general purpose technologies. Working Paper 25148, National Bureau of Economic Research.

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

Cong, W., T. Liang, and X. Zhang (2019). Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information.

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

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics 132*(4), 1593–1640.

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

Eisfeldt, A. L., A. Falato, and M. Z. Xiaolan (2021). Human capitalists. Working Paper 28815, National Bureau of Economic Research.

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

Field, A. J. (2003). The most technologically progressive decade of the century. *American Economic Review 93*(4), 1399–1413.

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

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

Goldschlag, N., T. J. Lybbert, and N. J. Zolas (2020). Tracking the technological composition of industries with algorithmic patent concordances. *Economics of Innovation and New Technology 29*(6), 582–602.

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

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

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *The RAND Journal of Economics 36*(1), 16–38.

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

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

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

Huckfeldt, C. (2021). Understanding the scarring effect of recessions. *American Economic Review forthcoming.*

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

Jaimovich, N. and H. Siu (2018). The trend is the cycle: Job polarization and jobless recoveries. *Review of Economics and Statistics, forthcoming.*

Jerome, H. (1934). *Mechanization in Industry.* NBER.

Jones, C. I. and J. Kim (2018). A schumpeterian model of top income inequality. *Journal of Political Economy 126*(5), 1785–1826.

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

Kambourov, G. and I. Manovskii (2009). Occupational specificity of human capital. *International Economic Review 50*(1), 63–115.

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

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2020). Measuring technological innovation over the long run. *American Economic Review: Insights forthcoming.*

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

Keynes, J. M. (1930). *Economic Possibilities For Our Grandchildren*, Volume 9 of *The Collected Writings of John Maynard Keynes*, pp. 321–332. Royal Economic Society.

Kogan, L., D. Papanikolaou, L. D. W. Schmidt, and J. Song (2020). Technological innovation and labor income risk. Working Paper 26964, National Bureau of Economic Research.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Kopytov, A., N. Roussanov, and M. Taschereau-Dumouchel (2018). Short-run pain, long-run gain? recessions and technological transformation. *Journal of Monetary Economics 97*, 29–44.

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

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

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

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics 13*(4), 653–677.

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

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

Webb, M. (2019). The impact of artificial intelligence on the labor market.

Zhang, M. B. (2019). Labor-technology substitution: Implications for asset pricing. *The Journal of Finance 74*(4), 1793–1839.

Figures and Tables

Figure 1: Examples of Technology Exposure: By Innovations

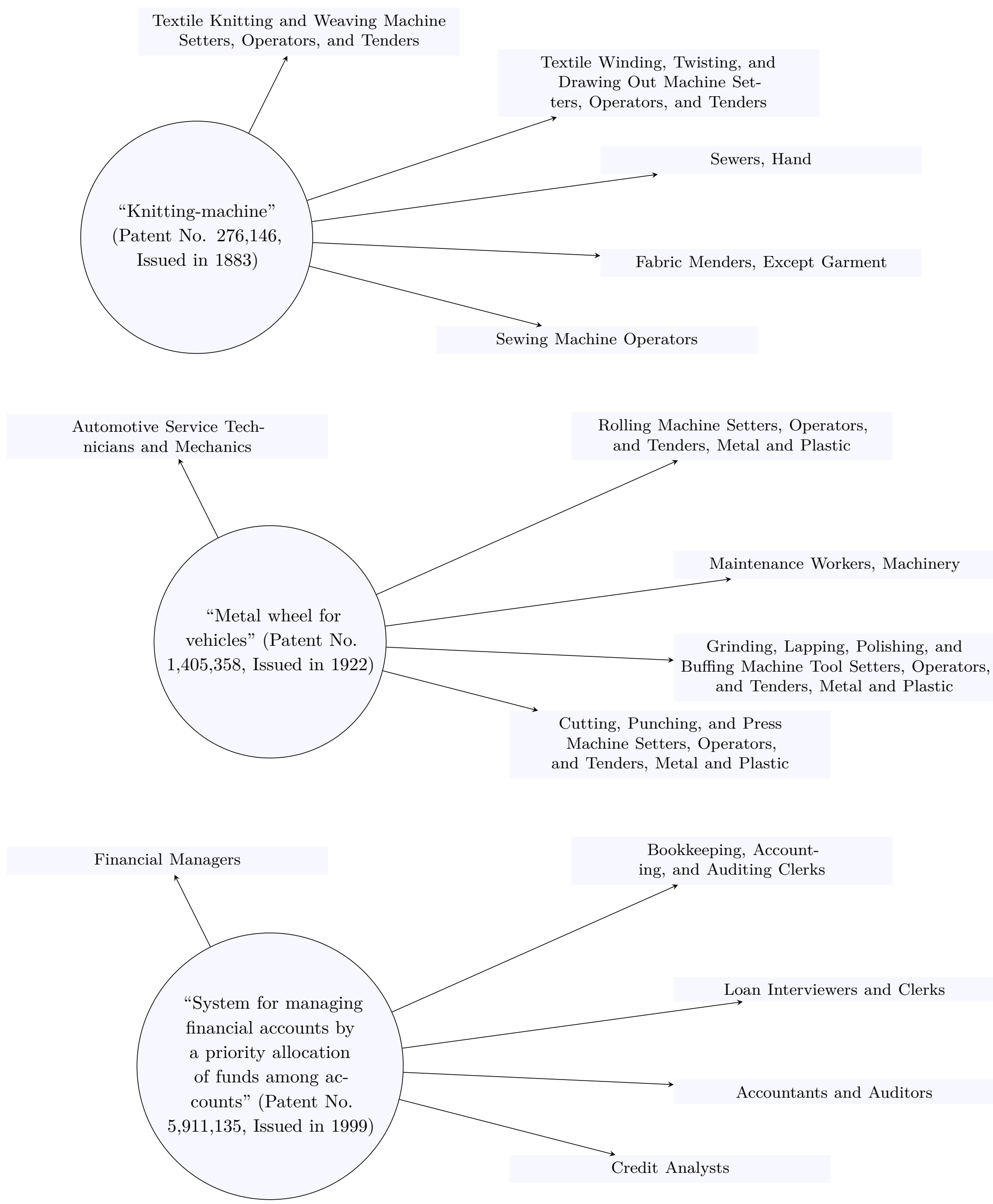
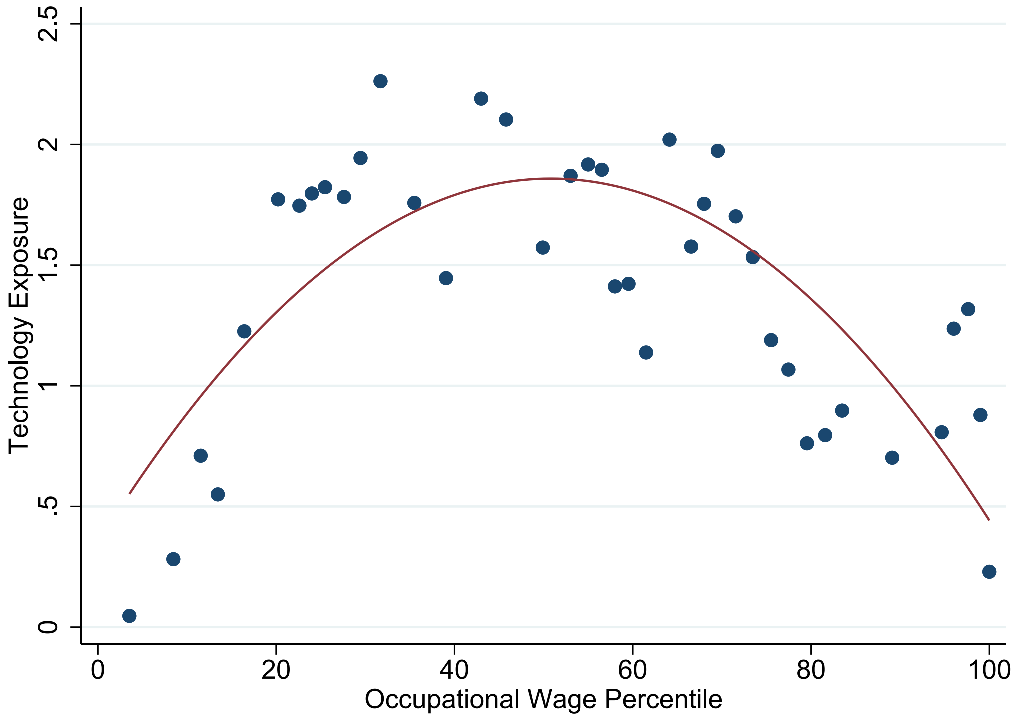
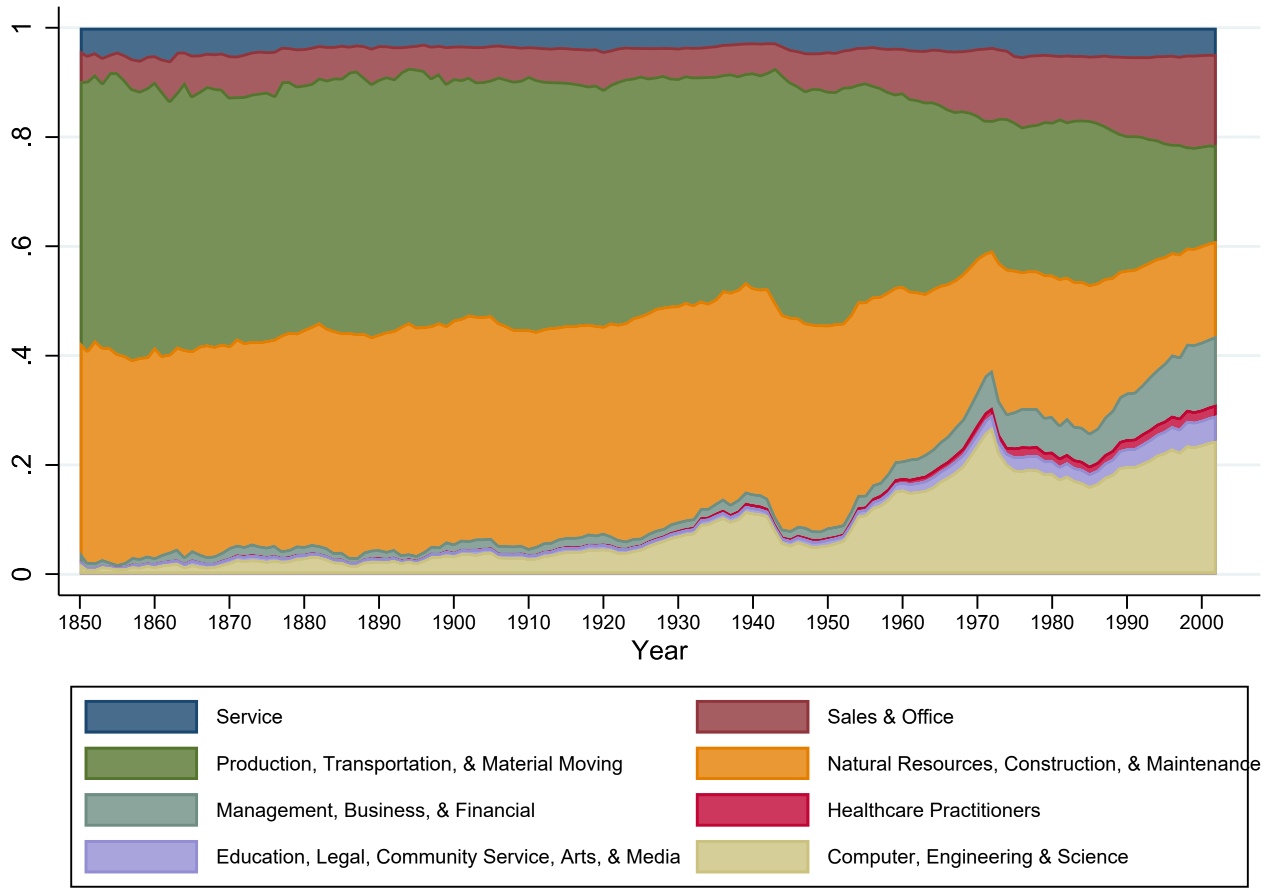


Figure 2: Technological Exposure, by occupation income level



Note: This figure plots average $\eta_{i,t}$ for occupations over the 1980 to 2002 period by wage percentile rank. The wage data come from the Current Population Survey Merged Outgoing Rotation Groups.

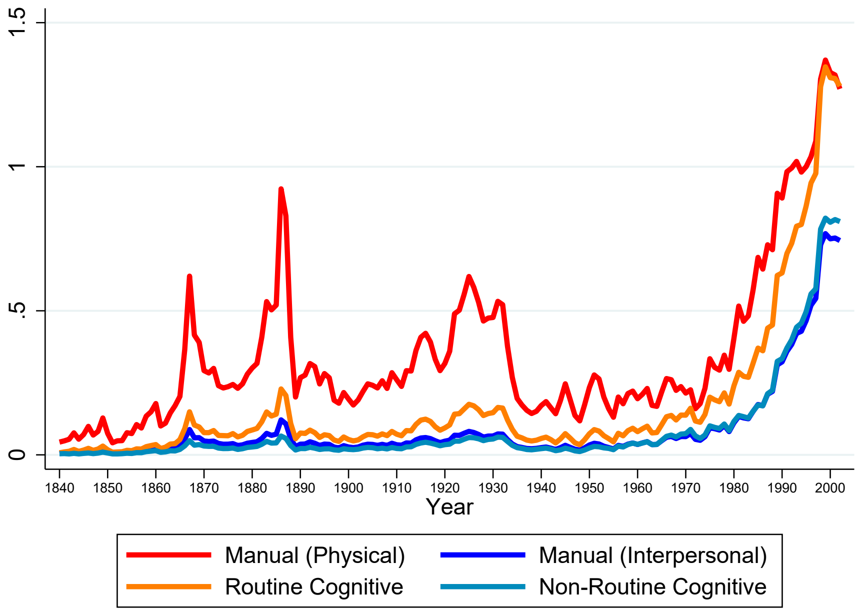
Figure 3: Technological Exposure, composition by major occupation group



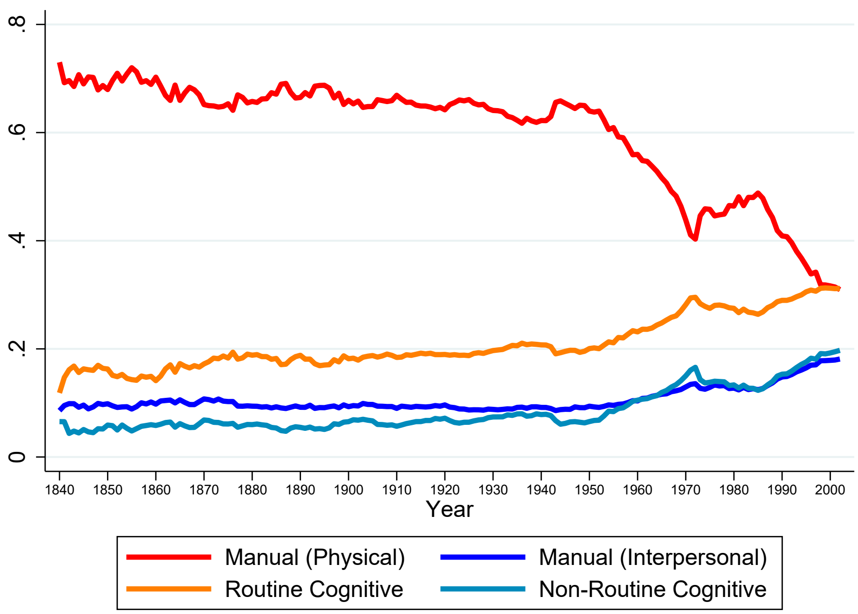
Note: This figure plots the average of our occupation-level innovation exposure index, $\eta_{i,t}$, where $\eta_{i,t}$ has been averaged separately within eight broad occupation groups. The occupation group averages are re-scaled each year so that the total across all groups sums to one in the given year.

Figure 4: Technology Exposure, by task type

A. Levels



B. Composition



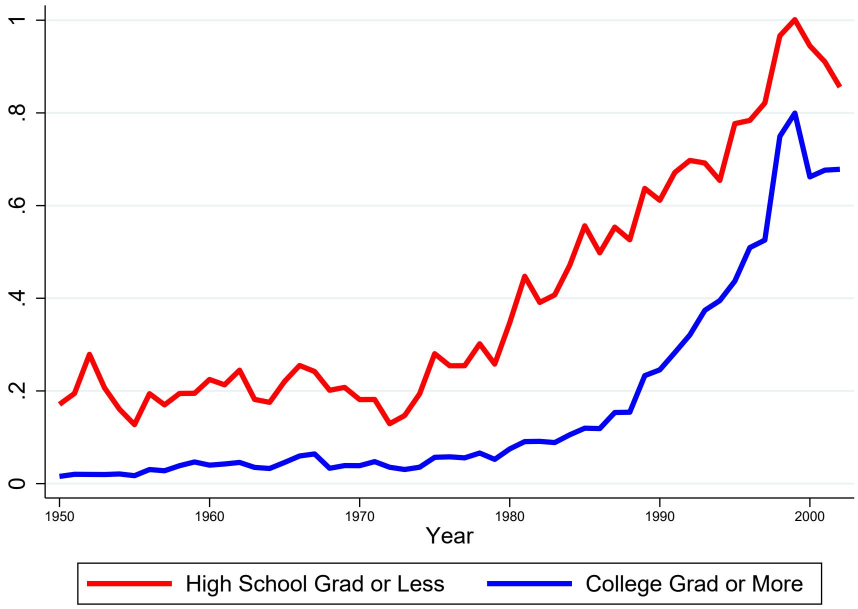
Note: This figure plots the level and composition of our index of technological exposure by task category:

$$\lambda_{w,t} = \sum_i \eta_{i,t} \times T_{w,t}(i) \times \omega_i \tag{35}$$

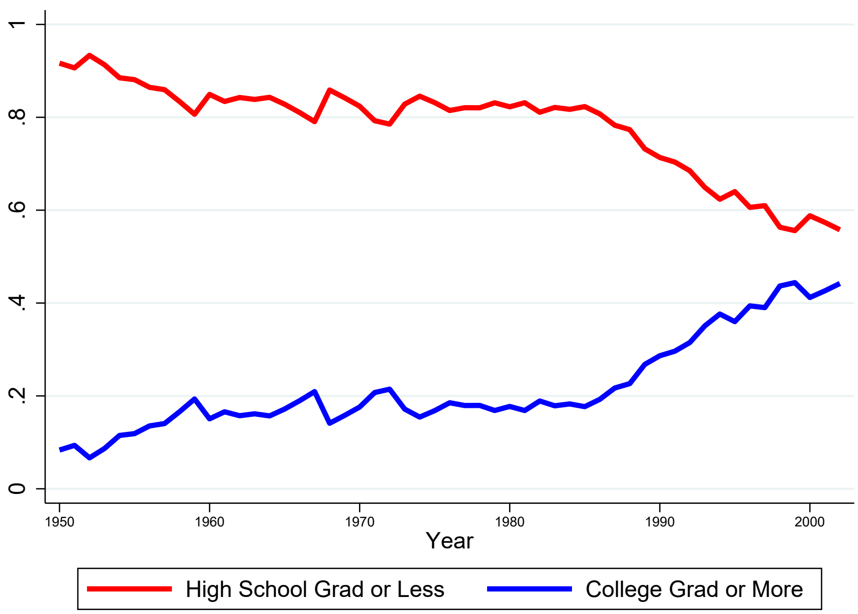
Panel A plots the raw index $\lambda_{w,t}$ and panel B plots the relative shares $\lambda_{w,t}/\sum_{w'} \lambda_{w',t}$ Here w represents one of the four given task categories. $T_{w,t}$ is an indicator that takes a value of 1 if occupation i is in the top quintile of the cross-sectional distribution of task scores for task category w. $\eta_{i,t}$ is our index of technological exposure and ω_i gives the Acemoglu and Autor (2011) occupational employment shares. See main text for more details.

Figure 5: Technology Exposure, by education requirements

A. Levels



B. Composition

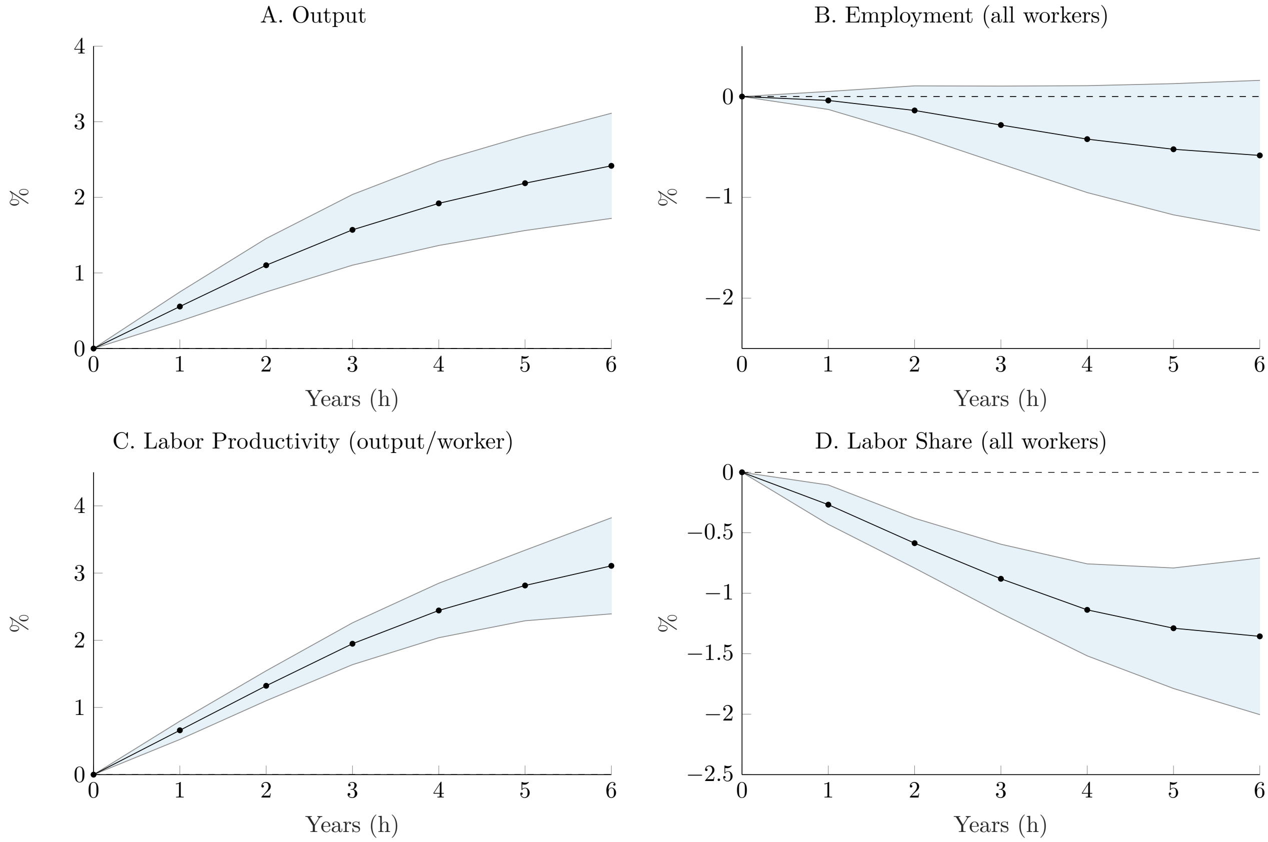


Note: This figure plots the level and composition of our index of technological exposure by education category:

$$\zeta_{s,t} = \sum_i \eta_{i,t} \times S_{s,t}(i) \times \omega_{i,t} \tag{36}$$

Panel A plots the raw index $\zeta_{s,t}$ and panel B plots the relative shares $\zeta_{s,t} / \sum_{s'} \zeta_{s',t}$ Here s represents either the educational category "high school or less" or "college grad or more". $S_{s,t}$ is and indicator that takes a value of 1 if occupation i is in the top quintile of the time t cross-sectional distribution of shares of workers falling in category s. $\eta_{i,t}$ is our index of technological exposure and $\omega_{i,t}$ gives occupational employment shares. See main text for more details.

Figure 6: Innovation: Productivity vs Labor Share

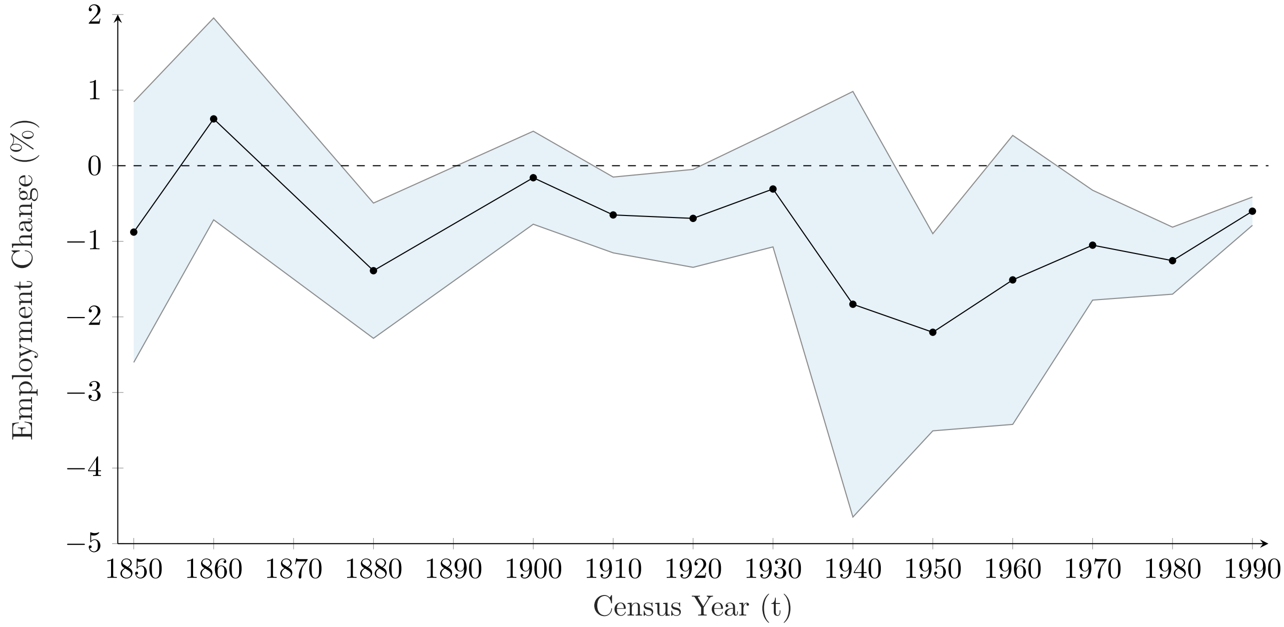


Note: The figure plots the estimated coefficients $\beta(k)$ from regressions of the form

$$\log X_{j,t+k} - \log X_{j,t} = \alpha(k) + \beta(k)\,\psi_{j,t} + \delta(k)Z_{j,t} + \epsilon_{j,t} \qquad \text{for } k = 1 \ldots T \text{ years}$$

The main independent variable $\psi_{j,t}$ is an index of innovation in industry j in year t, constructed as follows. First, we assign breakthrough patents to industries using the patent CPC tech class to industry crosswalk from Goldschlag et al. (2020). Second, we only include breakthrough patents whose average similarity to the industry's occupations (using occupation-by-industry employment weights) are above the (unconditional) median. We scale $\psi_{j,t}$ by US population and normalize to unit standard deviation. Controls $Z_{j,t}$ include industry employment shares, year fixed effects and lagged 5-year growth rate of the dependent variable. Standard errors are clustered by industry, and corresponding t-stats are shown in parentheses.

Figure 7: Employment and Technology Exposure (long-run: 1850–present)

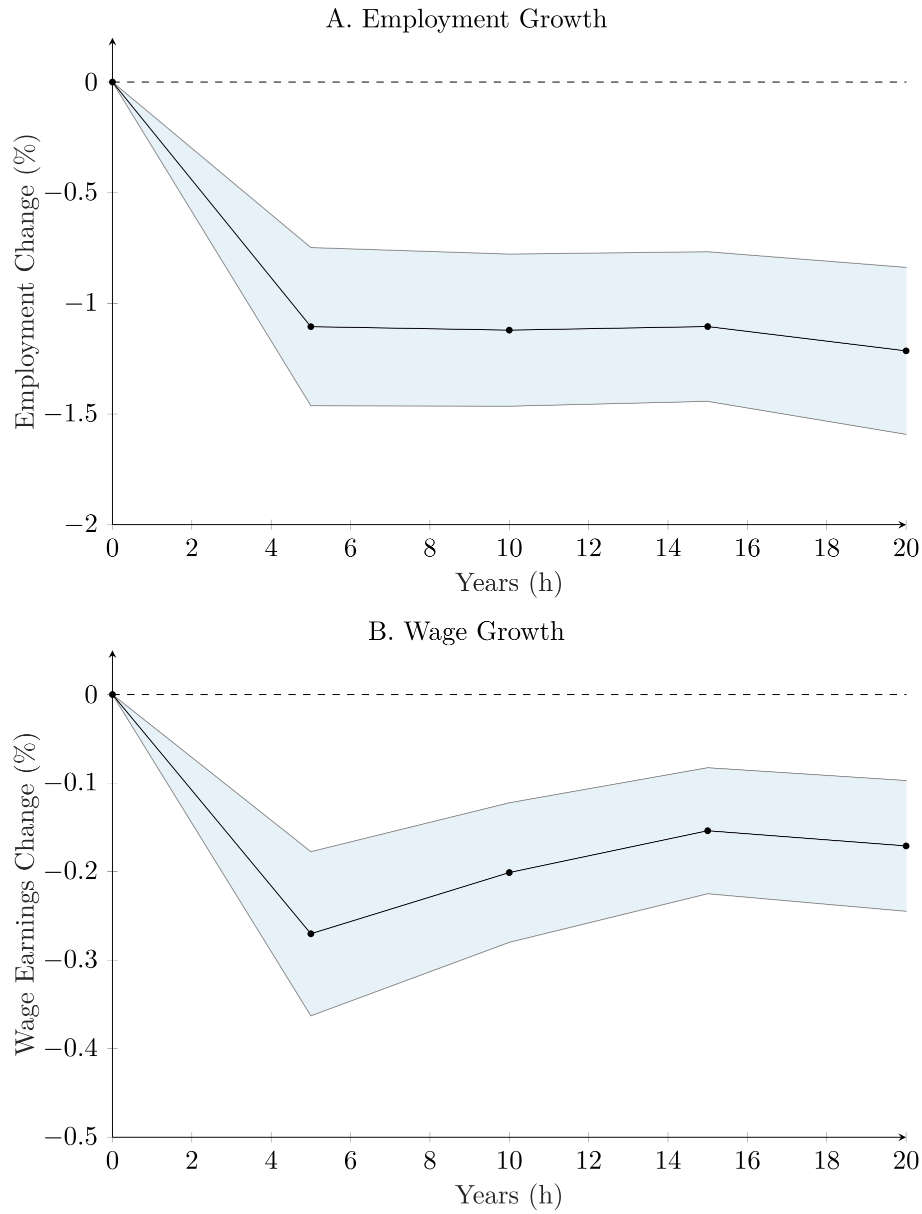


Note: Figure shows the slope coefficients on annual regressions of 20-Year employment share growth on our technology exposure $\eta_{i,t}$, using Census Years from 1850 to 1990. Specifically, we plot the β coefficients from

$$\frac{1}{k}\left(\log Y_{j,t+k} - \log Y_{j,t}\right) = \alpha_t + \beta_t \eta_{i,t} + \lambda_t + \epsilon_{i,t}$$

Here $Y_{i,t}$ is the occupation's share in total non-farm employment. Standard errors are clustered by occupation and shaded area represents the corresponding 90% confidence intervals for β_τ. Growth rates are expressed in annualized percentage terms and $\eta_{i,t}$ is standardized.

Figure 8: Employment, wage earnings and technology exposure (recent period: 1980–present)



Note: The Figures above plot coefficients from panel regressions of annualized wage and income growth rates over different time horizons on occupation innovation exposures:

$$y_{i,t+k} - y_{i,t} = \alpha + \beta \eta_{i,t} + \delta X_{i,t} + \varepsilon_{i,t}$$

Controls $X_{i,t}$–includes three one-year lags of dependent variable, and time fixed effects. Dependent variable is expressed in annualized percentage terms and $\eta_{i,t}$ is standardized. Figures plot 90% confidence interval for each time horizon. Data come from the CPS Merged Outgoing Rotation Groups (MORG) and cover the 1985–2018 period.

Figure 9: Employment share over the business cycle (Technology Exposure vs RTI)

Note: The above figure plots aggregate employment shares over time for occupations that were in the top quintiles of innovation exposure ($\eta_{i,t}$) and routine-task intensity in the year 1985. Vertical shaded bars represent NBER recession dates. Data source: CPS Merged Outgoing Rotation Groups extracts obtained from the Center For Economic Policy Research website.

Figure 10: Example: Order Clerks versus Personnel and Library Clerks

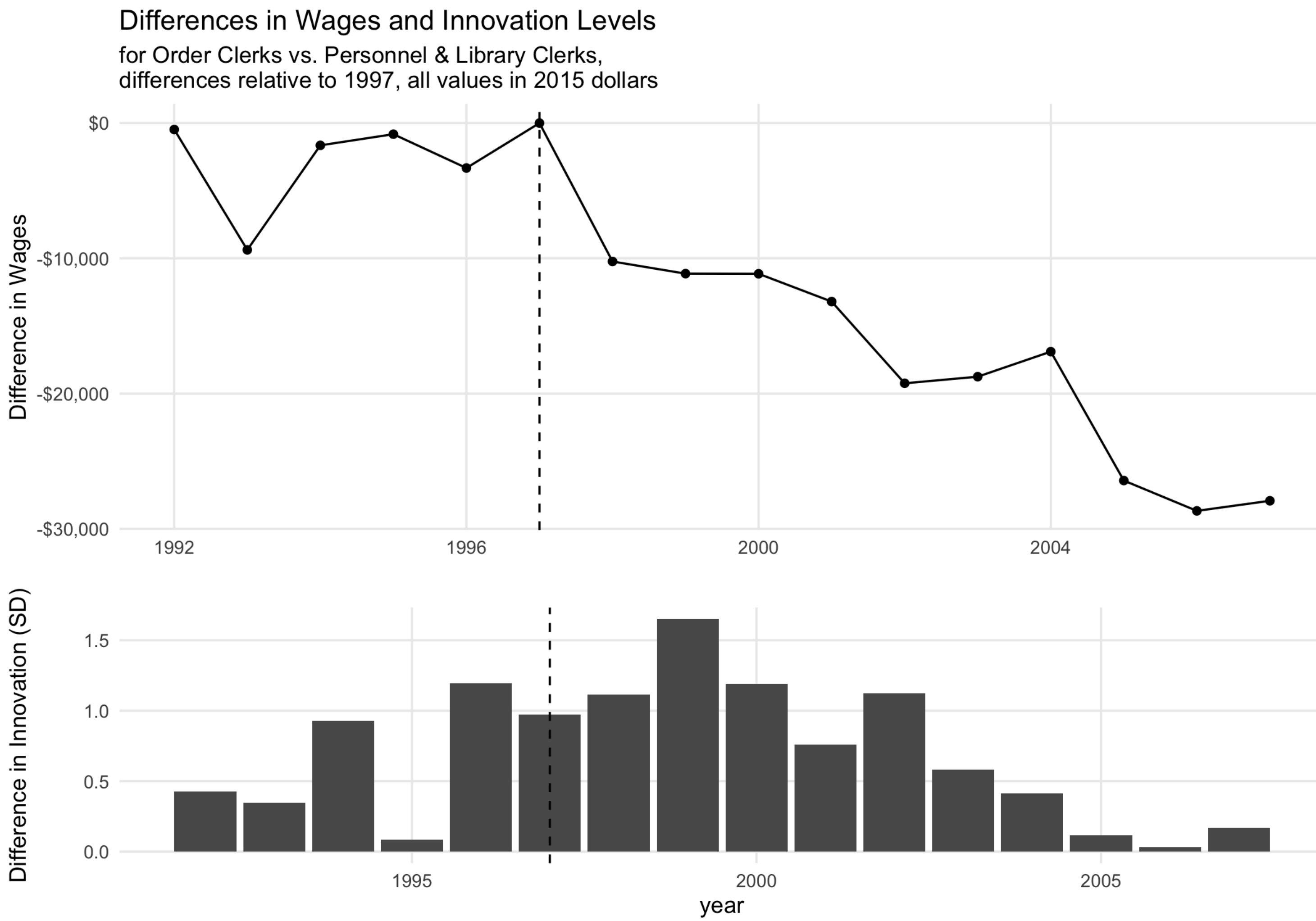
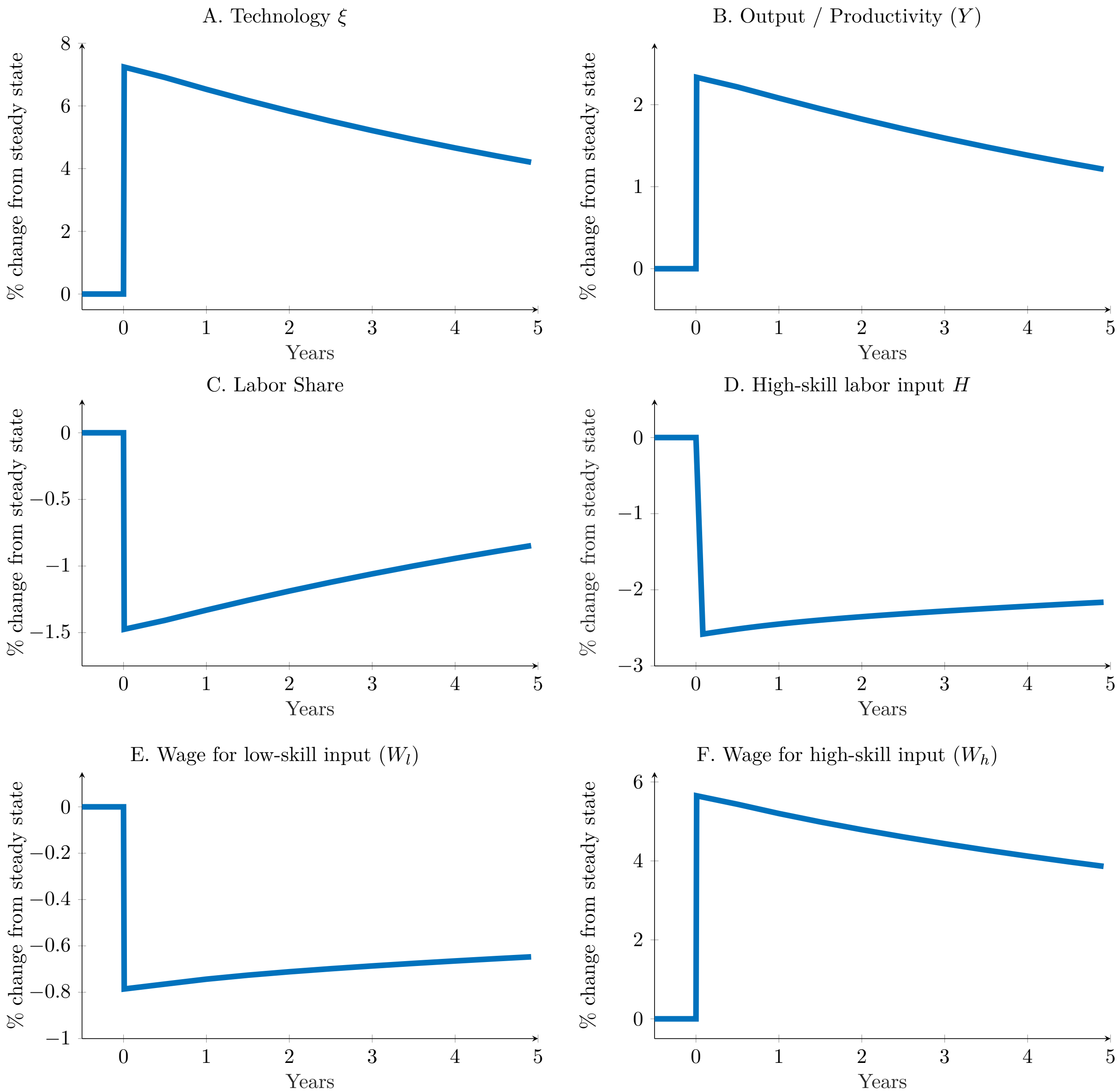


Figure 11: Model: Impulse Responses

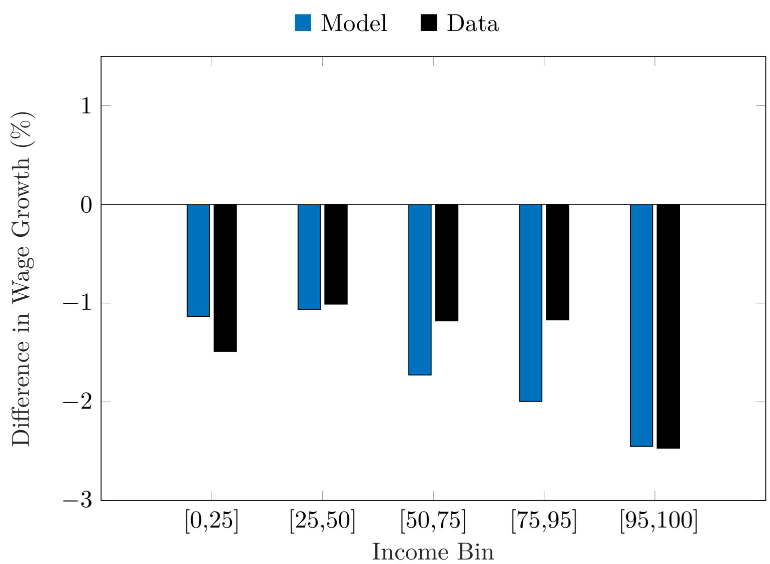


Note: This figure shows the impulse responses of key model quantities following a one-standard deviation technology shock evaluated at the steady state of the model.

Figure 12: Model: Innovation and Worker Earnings

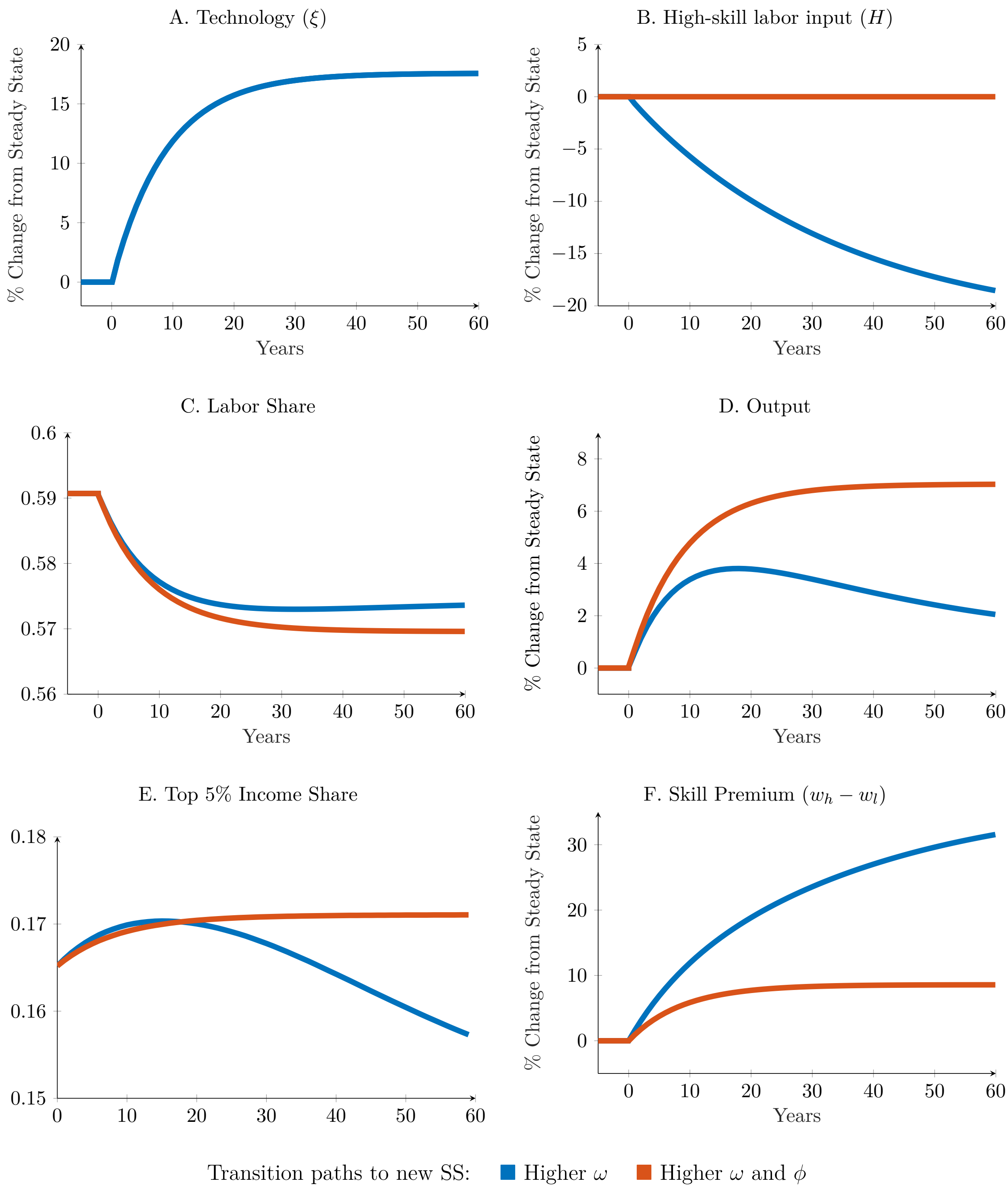
A. Differences in Post-Shock Wage Growth

B. Regression Coefficients for Post-Shock Wage Growth



Note: Panel A shows raw differences between wage growth during a shock period and wage growth if there had not been a shock for workers who are exposed to the shock, workers who are not exposed to the shock. Panel B shows the associated regression coefficients, which represent the marginal effects of a shock on wage growth given exposure. The left part of the figure shows the results in our baseline calibration. The right part of the figure compares to the case where there is no displacement of human capital.

Figure 13: The race between education and technology



Note: Figure computes the transition paths from the old to the new steady state for two permanent parameter shifts: 1) the blue line plots a permanent increase in the frequency of technological innovation ω, calibrated so that the level of technology ξ is permanently higher by one standard deviation relative to the old steady state (panel A); and 2) the orange line plots the transition paths associated with the same shift in ω but also with an increase in the rate of new skill acquisition ϕ such that the total supply of the high-skill labor input remains the same as the old steady state (panel B). Panel C plots the labor share of output; panel D plots total output/productivity; panel E plots income inequality, defined as the top 5% income share in the model; and panel F plots the skill premium.

Table 1: Technology And Employment Over the Long Run (1850–present)

	A. Occupation-level Employment				B. Industry X Occupation level employment			
	10 Years	20 Years	10 Years	20 Years	10 Years	20 Years	10 Years	20 Years
Technology Exposure, $\eta_{i,t}$	-0.43***	-0.75***	-0.33***	-0.66***	-0.37***	-0.76***	-0.38***	-0.86***
	(-4.68)	(-6.30)	(-4.17)	(-6.33)	(-2.76)	(-3.69)	(-2.83)	(-3.92)
Observations	2,865	2,574	2,492	2,208	102,400	81,009	72,451	54,662
R^2 (Within)	0.016	0.043	0.067	0.078	0.003	0.013	0.004	0.018
Controls								
Time FE	Y	Y	Y	Y				
Industry X Time FE					Y	Y	Y	Y
Lagged Dependent Variable			Y	Y			Y	Y

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

$$\frac{1}{k}\Big(\log Y_{i,t+k} - \log Y_{i,t}\Big) = \alpha_0 + \alpha_t + \beta(k)\eta_{i,t} + \rho\left(\log Y_{i,t} - \log Y_{i,t-k}\right) + \epsilon_{i,t}$$

for $k = 10, 20$ years for Census years spanning from 1850-2010. Here $Y_{i,t}$ is the occupation's share in total non-farm employment. $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. Standard errors are clustered by occupation and corresponding t-stats are shown in parentheses. Observations are weighted by occupation employment share at time t. Census year 1870 does not show up in the first column of the 20-year subsample regressions because the 1890 Census records no longer exist.

Table 2: Summary Statistics: Worker-Level Data

Variable	Mean	SD	5%	25%	Median	75%	95%
W2-Earnings	100,100	282,400	18,060	47,100	76,010	119,600	233,900
Age	41	7	29	34	41	47	53
Age: Lowest 25% Earners	39	7	29	33	38	44	51
Age: Top 5% Earners	44	7	31	39	46	50	54
Earnings growth, 3-years	-0.053	0.573	-0.890	-0.124	0.015	0.143	0.540
Earnings growth, 5-years	-0.063	0.592	-0.968	-0.158	0.011	0.156	0.574
Earnings growth, 10 -years	-0.073	0.639	-1.098	-0.215	0.005	0.190	0.666

Note: The table reports summary statistics for our wage earnings data from the Census-CPS sample, which covers the 1988 to 2016 period. W2-Earnings are reported in terms of 2015 dollars.

Table 3: Worker Earnings and Technology Exposure

	(1)	(2)
A. Cond. Mean: $E[g]$, by Horizon		
3 years	-1.21	-1.12
	(-6.32)	(-5.65)
5 years	-1.55	-1.30
	(-6.84)	(-4.77)
10 years	-1.43	-1.24
	(-4.96)	(-3.14)
B. Risk: Absolute Income Growth $E[\|g\|]$		
3 years	0.73	0.25
	(3.03)	(1.59)
5 years	0.76	0.37
	(2.40)	(1.78)
10 years	0.38	0.47
	(0.95)	(1.68)
C. Skewness: Prob. Large Income Decline $p(g < p^{10})$		
3 years	0.56	0.38
	(6.31)	(4.18)
5 years	0.58	0.41
	(4.76)	(3.31)
10 years	0.37	0.27
	(2.09)	(1.48)
Controls:		
Industry FE	Y	
Occupation FE	Y	
Year FE	Y	
Industry × Year FE		Y
Occupation × Year FE		Y

Note: Panel A shows the estimated slope coefficients $\beta(h)$ (times 100) from equation (21) in the main text for horizons h of 3,5, and 10 years. Panels B and C focus on the 5-year horizon. Panel B shows the slope coefficients of a variant of the above specification where we replace the dependent variable $g_{i,t:t+h}$ with its absolute value $|g_{i,t:t+h}|$ to capture the response of second moments to changes in technology exposure $\eta_{i,t}$. Similarly, Panel C replaces the dependent variable with a dummy that takes the value of one if $g_{i,t:t+h}$ lies in the bottom 10-th percentile; this specification allows us to capture increases in negative skewness in response to an increase in $\eta_{i,t}$. We report t-statistics (in parentheses) using standard errors clustered at the industry (NAICS 4-digit) level. All specifications include industry times year and occupation times year fixed effects. We normalize $\eta_{i,t}$ to unit standard deviation. The bottom panel shows the p-values associated with the hypotheses that the coefficients are equal across the reported subgroups.

Table 4: Worker Earnings and Technology Exposure, by Education

Education	(1)	(2)	(3)	(4)	(5)
	Cond. Mean $E[g]$			St. Dev $E[\|g\|]$	Skew $p(g < p^{10})$
	3-year	5-year	10-year	5-year	5-year
College	-0.91	-1.14	-1.21	0.49	0.35
	(-3.33)	(-3.34)	(-2.49)	(2.65)	(2.57)
No College	-1.30	-1.50	-1.76	0.39	0.48
	(-5.58)	(-5.18)	(-4.47)	(1.51)	(4.33)
Coeff. Differences			p-values		
College = No College	0.043	0.110	0.052	0.537	0.071

Note: Columns (1) to (3) show the estimated slope coefficients (times 100) from equation (21) in the main text: the dependent variable is worker earnings growth over horizons of 3,5, and 10 years; the main independent variable of interest is a worker's technology exposure $\eta_{i,t}$. The slope coefficient $\beta(h)$ is allowed to vary with the worker's education. Columns (1) to (3) correspond to horizons of 3,5, and 10 years. Columns (4) and (5) focus on the 5-year horizon. Column (4) shows the slope coefficients of a variant of the above specification where we replace the dependent variable $g_{i,t:t+h}$ with its absolute value $|g_{i,t:t+h}|$ to capture the response of second moments to changes in technology exposure $\eta_{i,t}$. Similarly, Column (5) replaces the dependent variable with a dummy that takes the value of one if $g_{i,t:t+h}$ lies in the bottom 10-th percentile; this specification allows us to capture increases in negative skewness in response to an increase in $\eta_{i,t}$. We report t-statistics (in parentheses) using standard errors clustered at the industry (NAICS 4-digit) level. All specifications include industry times year and occupation times year fixed effects.We normalize $\eta_{i,t}$ to unit standard deviation. The bottom panel shows the p-values associated with the hypotheses that the coefficients are equal across the reported subgroups.

Table 5: Worker Earnings and Technology Exposure, by Prior Income

	(1)	(2)	(3)	(4)	(5)
Income Percentile	Cond. Mean $E[g]$			St. Dev $E[\|g\|]$	Skew $p(g < p^{10})$
	3-year	5-year	10-year	5-year	5-year
0 to 25-th	-1.24	-1.49	-1.85	-0.26	0.29
	(-5.76)	(-5.28)	(-4.70)	(-1.07)	(2.23)
25 to 50-th	-0.85	-1.01	-0.96	0.10	0.26
	(-4.14)	(-3.57)	(-2.04)	(0.40)	(1.45)
50 to 75-th	-1.01	-1.18	-1.01	0.48	0.39
	(-3.51)	(-3.07)	(-1.87)	(1.65)	(2.84)
75 to 95-th	-1.05	-1.17	-0.81	0.76	0.39
	(-4.12)	(-3.45)	(-1.64)	(3.25)	(2.76)
95 to 100-th	-2.24	-2.47	-2.28	2.01	1.26
	(-6.11)	(-4.75)	(-3.83)	(5.78)	(5.50)
Coeff. Differences	p-values				
[95-100] = [25-95]	0.000	0.000	0.002	0.000	0.000
[0-25] = [25-95]	0.610	0.630	0.331	0.175	0.853
[95-100] = [0-25]	0.019	0.092	0.498	0.000	0.000

Note: Columns (1) to (3) show the estimated slope coefficients (times 100) from equation (21) in the main text: the dependent variable is worker earnings growth over horizons of 3,5, and 10 years; the main independent variable of interest is a worker's technology exposure $\eta_{i,t}$. The slope coefficient $\beta(h)$ is allowed to vary with the worker's prior income rank. Columns (1) to (3) correspond to horizons of 3,5, and 10 years. Columns (4) and (5) focus on the 5-year horizon. Column (4) shows the slope coefficients of a variant of the above specification where we replace the dependent variable $g_{i,t:t+h}$ with its absolute value $|g_{i,t:t+h}|$ to capture the response of second moments to changes in technology exposure $\eta_{i,t}$. Similarly, Column (5) replaces the dependent variable with a dummy that takes the value of one if $g_{i,t:t+h}$ lies in the bottom 10-th percentile; this specification allows us to capture increases in negative skewness in response to an increase in $\eta_{i,t}$. We report t-statistics (in parentheses) using standard errors clustered at the industry (NAICS 4-digit) level. All specifications include industry times year and occupation times year fixed effects. We normalize $\eta_{i,t}$ to unit standard deviation. The bottom panel shows the p-values associated with the hypotheses that the coefficients are equal across the reported subgroups.

Table 6: Worker Earnings and Technology Exposure, by Age

Worker Age	(1)	(2)	(3)	(4)	(5)
	Cond. Mean $E[g]$			St. Dev $E[\lvert g \rvert]$	Skew $p(g < p^{10})$
	3-year	5-year	10-year	5-year	5-year
25–35 years	-0.39	-0.64	-0.92	0.38	0.18
	(-1.86)	(-2.55)	(-2.42)	(1.84)	(1.50)
35–45 years	-0.86	-1.04	-1.37	0.05	0.23
	(-5.24)	(-5.08)	(-3.63)	(0.37)	(2.55)
45–55 years	-1.95	-2.25	-2.51	1.13	0.88
	(-3.72)	(-3.55)	(-3.02)	(2.5)	(3.36)
Coeff. Differences			p-values		
45–55 = 25–35	0.001	0.002	0.006	0.051	0.001
45–55 = 35–45	0.015	0.020	0.044	0.010	0.011

Note: Columns (1) to (3) show the estimated slope coefficients (times 100) from equation (21) in the main text: the dependent variable is worker earnings growth over horizons of 3,5, and 10 years; the main independent variable of interest is a worker's technology exposure $\eta_{i,t}$. The slope coefficient $\beta(h)$ is allowed to vary with the worker's age. Columns (1) to (3) correspond to horizons of 3,5, and 10 years. Columns (4) and (5) focus on the 5-year horizon. Column (4) shows the slope coefficients of a variant of the above specification where we replace the dependent variable $g_{i,t:t+h}$ with its absolute value $|g_{i,t:t+h}|$ to capture the response of second moments to changes in technology exposure $\eta_{i,t}$. Similarly, Column (5) replaces the dependent variable with a dummy that takes the value of one if $g_{i,t:t+h}$ lies in the bottom 10-th percentile; this specification allows us to capture increases in negative skewness in response to an increase in $\eta_{i,t}$. We report t-statistics (in parentheses) using standard errors clustered at the industry (NAICS 4-digit) level. All specifications include industry times year and occupation times year fixed effects.We normalize $\eta_{i,t}$ to unit standard deviation. The bottom panel shows the p-values associated with the hypotheses that the coefficients are equal across the reported subgroups.

Table 7: Model Parameters

Description	Parameter	Value
Share of workers who do not move up the ladder	s_l	0.375
Minimum level of skill	$\underline{\theta}$	0.03
Probability of worker exit	δ	0.025
Amount of skills acquired	m	0.03
CES parameter in inner nest (technology ξ and low-skill labor L)	ρ	0.74
Share of technology in inner nest	λ	0.27
CES parameter in outer nest (high-skill labor H and ξ/L composite)	σ	-0.12
Share of high-skill labor in outer nest	μ	0.17
Size of technology improvement	κ	0.31
Arrival rate of technology shocks	ω	1.56
Share of exposed workers	α	0.32
Human capital loss percentage conditional on fall	h	0.06
Rate of depreciation of technology	g	0.12
Likelihood of worker skill acquisition	ϕ	2.40

Note: Table reports the parameter used to calibrate the model. The first four parameters are calibrated a priori; the latter 10 parameters are chosen to fit the statistics reported in Table 8.

Table 8: Model Fit

Statistic	Data	Model
Labor share, average	0.66	0.59
Labor share, response to ξ	-1.29	-1.48
Skill premium (p75 / p25 ratio), average	2.45	1.68
Labor productivity, response to ξ	2.81	2.31
Worker earnings growth response to ξ		
0 to 25-th percentile	-1.49	-1.13
25 to 50-th percentile	-1.01	-1.06
50 to 75-th percentile	-1.18	-1.72
75 to 95-th percentile	-1.17	-2.00
95 to 100-th percentile	-2.47	-2.45
Likelihood of large wage declines in response to ξ		
0 to 25-th percentile	0.29	0.51
25 to 50-th percentile	0.26	0.51
50 to 75-th percentile	0.39	0.51
75 to 95-th percentile	0.39	0.51
95 to 100-th percentile	1.26	1.40

Note: Table reports the fit of the model to the statistics that we target. The parameters used in our calibration are listed in Table 8.

Table 9: Technology and Labor Market Outcomes: Comparison to Other Measures

	A. Employment				
	(1)	(2)	(3)	(4)	(5)
Technology Exposure $\eta_{i,1980}$	-0.79***	-0.87***	-0.74***	-0.74***	-0.82***
	(-5.68)	(-6.25)	(-5.26)	(-5.12)	(-5.99)
Routine Task Intensity (RTI)		-0.058			0.011
		(-0.50)			(0.08)
Offshorability		-0.15**			-0.20**
		(-2.04)			(-2.50)
Robot Exposure			-0.66**		-0.93**
			(-2.21)		(-2.21)
Software Exposure				-0.36	0.069
				(-1.34)	(0.21)
	B. Wages				
	(1)	(2)	(3)	(4)	(5)
Technology Exposure $\eta_{i,j,1980}$	-0.082***	-0.064***	-0.071***	-0.100***	-0.078***
	(-4.11)	(-3.11)	(-3.87)	(-5.37)	(-4.21)
Routine Task Intensity (RTI)		-0.066*			-0.059*
		(-1.97)			(-1.70)
Offshorability		0.0042			-0.0026
		(0.17)			(-0.09)
Robot Exposure			-0.13*		-0.27***
			(-1.86)		(-2.76)
Software Exposure				0.13**	0.24***
				(2.10)	(4.73)
Observations	17536	16959	17448	17448	16959

Note: This table shows results from estimating

$$\frac{1}{k}\Big(\log Y_{i,j,t+k} - \log Y_{i,j,t}\Big) = \alpha + \alpha_j + \beta\eta_{i,1980} + \delta X_i + \epsilon_{i,j}$$

Here i indexes occupations and j indexes industries; we report results for $k = 32$ years using Deming (2017) data from the 1980 Census and the 2012 ACS. The dependent variables are employment (Panel A) or average wages (Panel B). All specifications include industry fixed effects and controls for occupation employment share in 1980, occupation log wage in 1980, three categorical indicators for the occupation's average education level in 1980. We additionally include the routine-task intensity and the measure of occupation-level offshorability from Autor and Dorn (2013) and the measure of exposure to robots or software from Webb (2019) depending on the specification. Observations are weighted by employment share in 1980.

A Appendix

A.1 Converting Patent Text for Numerical Analysis

Here, we briefly overview our conversion of unstructured patent text data into a numerical format suitable for statistical analysis. We obtain text data for measuring patent/job task similarity from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. We use the patent text data parsed from the USPTO patent search website in Kelly et al. (2020), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET. We then combine all tasks for a given occupation at the 2010 O*NET 6-digit level into one occupation-level corpus. The process for cleaning and preparing the text files for numerical representation follows the steps outlined below.

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

Stop words come from the following sources:

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

- `http://www.nltk.org/book/ch02.html` (NLTK stop words list)

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

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

A.2 Description of Word Embedding Vectors

To appreciate how our metric differs from the standard bag-of words approach it is useful to briefly examine how word embeddings are computed in Pennington et al. (2014). Denote the matrix X as a $V \times V$ matrix of word co-occurence counts obtained over a set of training documents, where V is the number of words in the vocabulary. Then $X_{i,j}$ tabulates the number of times word j appears in the context of the word i.[19] Denote $X_i = \sum_k X_{i,k}$ as the number of times any word appears in the context of word i, and the probability of word j occuring in the context of word i is $P_{i,j} \equiv X_{i,j}/X_i$. The goal of the word embedding approach is to construct a mapping $F(\cdot)$ from some d-dimensional vectors x_i, x_j, and $\tilde{x}_k$ such that

$$F(x_i, x_j, \tilde{x}_k) = \frac{P_{i,k}}{P_{j,k}} \tag{A.1}$$

Imposing some conditions on the mapping $F(\cdot)$, they show that a natural choice for modeling $P_{i,k}$ in (A.1) is

$$x_i^T \tilde{x}_k = \log(X_{i,k}) - \log(X_i) \tag{A.2}$$

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

$$x_i^T \tilde{x}_k + b_i + b_k = \log(X_{i,k}) \tag{A.3}$$

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

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

$$\text{Min}_{x_i,\tilde{x}_k,b_i,b_k} \sum_{i=1}^{V}\sum_{j=1}^{V} f(X_{i,j}) \left(x_i^T \tilde{x}_k + b_i + b_k - \log(X_{i,j})\right)^2 \tag{A.4}$$

Here the observation-specific weighting function $f(X_{i,j})$ equals zero for $X_{i,j} = 0$ so that the log is well defined, and is constructed to avoid overweighting rare occurences or extremely frequent occurences. The objective (A.4) is a highly-overidentified least squares minimization problem. Since the solution is not unique, the model is trained by randomly instantiating x_i and $\tilde{x}_k$ and performing gradient descent for a pre-specified number of iterations, yielding d-dimensional vector representations of a given word. Here d is a hyper-parameter; Pennington et al. (2014) find that $d = 300$ works well on word analogy tasks.

Since (A.4) is symmetric it yields two vectors for word i, x_i and $\tilde{x}_i$, so the final word vector is taken as the average of the two. The ultimate output is a dense 300-dimensional vector for each word i that has been estimated from co-occurence probabilities and occupies a position in a word vector space such that the pairwise distances between words (i.e. using a metric like the cosine similarity) are related to the probability that the words occur within the context of one another and within the context of other similar words. Note that the basis for this word vector space is arbitrary and has no meaning; distances between word embeddings are only well-defined in relation to one another and a different training instance of the same data would yield different word vectors but very similar pairwise distances between word vectors.

Our method for backing out a geometric representation of the "meaning" of a document in (7) is to construct a weighted average of the meaning of all words in the document. Thus our vector representation of documents retains the 300-dimensional structure of the individual word constituents; these vectors are much denser and smaller than the very large and sparse document vectors in the standard bag of words methodology. In brief, there are two key characteristics that differentiate our approach relative to bag of words techniques. First, X_i is no longer a sparse vector like V_i. Moreover, because of the way word vectors are estimated, our method allows vectors containing similar words to be "close" to one another. Thus, relative to the bag of words approach our method: (1) constitutes a large dimensionality reduction; and, (2) can incorporate a notion of synonyms/distances between word meanings.

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

$$\text{Sim}_{i,j} = \frac{X_i}{||X_i||} \cdot \frac{X_j}{||X_j||} \tag{A.5}$$

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

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

$$V_1 = [1, 1, 0, 0], \quad \text{and} \quad V_2 = [0, 0, 1, 1] \tag{A.6}$$

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

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

$$X_1 = (1/2) \times \log(2) x_{dog} + (1/2) \times \log(2) x_{cat} \tag{A.7}$$

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

Last, our methodology bears some similarities to recent work by Webb (2019), who also analyzes the similarity between a patent and O*NET job tasks. Webb (2019) focuses on similarity in verb-object pairs in the title and the abstract of patents with verb-object pairs in the job task descriptions and restricts his attention to patents identified as being related to robots, AI, or software. He uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings. By contrast, we infer document similarity by using geometric representations of word meanings (GloVe) that have been estimated directly from word co-occurence counts. In addition to employing a different methodology, we also have a broader focus: we compute occupation-patent distance measures for all occupations and the entire set of USPTO patents since 1836. Furthermore, we use not only the abstract but the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention.

A.3 Constructing the Industry Innovation Measure

For each breakthrough patent p we assign it to an industry using probabilistic patent CPC tech class to NAICS crosswalks constructed by Goldschlag et al. (2020). The Goldschlag et al. (2020) crosswalk assigns probabilities that patents from a given technology class originated from a particular NAICS industry for different levels of NAICS aggregation. The NBER manufacturing database reports data the 6-digit NAICS code level, and so we use the Goldschlag et al. (2020) 6-digit NAICS to 3-digit CPC probabilistic crosswalk. We then aggregate the data to the 4-digit NAICS level to parallel the level of industry classification we use in our analysis in 3.2 of the main text.

Label the set of breakthrough patents issued in year t by Γ_t; $\alpha_{j,p}$ the probability of breakthrough patent p being issued to industry j, and κ_t the US population in year t. We then define the industry-level breakthrough patent index (including only patents with high average textual similarity to the industry workforce) by

$$\psi_{j,t} = \frac{1}{\kappa_t} \sum_{p \in \Gamma_t} \alpha_{j,p}) \tag{A.8}$$

A.4 Census public-use data

We gather Census data from IPUMS and compute aggregate employment shares for occupations in Census years spanning 1850-2010. We use the 1950 Census occupation definition for pre-1950 Census years since the more updated 1990 Census classification scheme is only available in post-1950 Census years. We make use of the 1990 Census occupation classifications for the years they are available. We then crosswalk Census occupations to the David Dorn occ1990dd classification scheme using the crosswalk files provided on his website and aggregate our measure $\eta_{i,t}$ to the occ1990dd-level by averaging across 6-digit SOC codes within an occ1990dd code. This results in a Census-year by occ1990dd panel of occupation employment shares. Census records for the year 1890 were destroyed in a fire, and so the employment growth observations for the 20-year horizon in 1870 or for the 10-year horizon in 1880 are not available.

For the post-1980 results, we use the Current Population Survey Merged Outgoing Rotation Groups (MORG). We obtain the cleaned versions of MORG extracts provided by the Center for Economic Policy Research (CEPR). We use the "wage3" variable that combines the usual hourly earnings for hourly workers and non-hourly workers, which adjusts for top-coding using a lognormal imputation and is constructed to match the NBER's recommendation for the most consistent hourly wage series from 1979 to the present. Using these data we construct a time series of wage and employment growth for occupations at the occ1990dd level. Because occ1990dd cannot be crosswalked to a balanced panel of occupations using the Census 1970 occupation codes, we start our analysis in the post-1982 time period when these extracts began using the 1980 Census occupation classification scheme.

A.5 Census-CPS administrative data

We use a random sample of individual workers tracked by the Current Population Survey (CPS) and their associated Detailed Earnings Records from the Census—which contains their W2 tax income. We limit the sample to individuals who are older than 25 and younger than 55 years old.

The CPS includes information on demographic information such as age and gender, but more importantly occupation at the time of the interview. We assign workers to occupations based on their response to the CPS survey (CPS "occ" variable). We construct a crosswalk between the yearly CPS occupations codes and the occ1990dd classification scheme and assign all CPS occupations their corresponding occ1990dd code. We assign this occupation to the worker for the next 5 years, thus effectively dropping observations where the CPS interview date is older than 5 years—so that the occupation information is relatively recent.

We merge the individual worker records from the Census-CPS matched sample to patent data at the industry (NAICS 4) level. Specifically, we identify the industry of where the patent origination by relying on the Census SSEL patent–assignee database, which provides a corresponding SSEL firm identifier ("firmid"), which we then use to obtain the firms' 4-digit NAICS code. In particular, we use two SSEL patent–assignee crosswalks: the newer Business Dynamics Statistics of Patenting Firms database (BDS-PF) and an older patent-SSL crosswalk created by Kerr and Fu (2008). The BDS-PF links are available starting with the 2000 SSL. We use the BDS-PF firmid-patent links for any patents for which it is available. Otherwise we rely on the links from Kerr and Fu (2008) created from the 1999 SSL. In cases where a firmid matches to multiple NAICS codes we apply the 4-digit NAICS code of highest employment. We drop any industry-year observations with no patents filed in a given year.

To allow the effects to vary with prior income, we assign workers into five groups based on their average income over the last three years (the last term in (18)) compared to workers in the same occupation and NAICS4 industry. These groups are defined based on the following percentiles of prior income [0%, 25%), [25%, 50%), [50%, 75%), [75%, 95%), [95%, 100%] calculated within industry–occupation cells. In the (uncommon) case when NAICS4 industries have cells which are too small to rank, we broaden the industry definition from 4 digit NAICS to 2 digit NAICS. Subsequently, any Industry–Occupation cells with fewer than 10 individuals are dropped.

A.6 Constructing a Statistical Displacement Factor

To construct our predictor we use a method proposed by Cong et al. (2019), which is well-suited to prediction exercises using large-scale textual data. Our adaptation of their method for the task of predicting occupation outcomes can be summarized in the following steps. Let the number of patent documents be N_p (where we restrict just to the set of breakthrough patents from Kelly et al. (2020) as described in section 2), the number of occupation task descriptions be N_o, and the number of words in the vocabulary formed from the union of all patent and occupation documents be N_w:

1. Perform approximate nearest neighbor search using a locality-sensitive hashing routine (LSH)

on vector representations of word meanings to form K clusters ("topics") of related words. Label the kth cluster of words C_k.

2. Create a $N_p \times N_w$ matrix of breakthrough patent documents by word counts weighted by term-frequency inverse document frequency (TF-IDF), computed over all patents (i.e. TF-IDF is computed also including non-breakthrough patents). Call this matrix A. Loop over each word cluster C_k from step 1 for $k = 1, \ldots, K$, and extract the submatrix of A formed by taking the columns in A corresponding to the words contained in cluster C_k. Call this submatrix A_k. Perform a singular-value decomposition of A_k and take its top singular value v_k (in absolute value) and corresponding top right singular vector V_k. Then take the $N_p \times 1$ vector $P_k = \frac{|A_k v_k|}{v_k' v_k}$ to be the loadings of each patent document on topic/word cluster k. Retain only the clusters C_k which rank in the top 500 based on their top absolute singular values.

3. Perform step 2 for all occupations, except only for the top 500 clusters that were retained. Call the resulting $N_o \times 1$ vector of occupation loadings O_k. Denote the set breakthrough of patents issued in year t by $\widehat{\Gamma}_t$. Let $O_{k,i}$ represent the ith element of O_k and $P_{k,j}$ the jth element of P_k, the vector of patent loadings on cluster k. Then occupation i's exposure to the kth topic in year t is given by

$$\psi_{i,k,t} = \frac{O_{k,i}}{\kappa_t} \sum_{j \in \widehat{\Gamma}_t} P_{k,j} \tag{A.9}$$

As before we only sum over breakthrough patents and normalize by U.S. population in year t (denoted by κ_t). This yields an occupation's exposure in each year to the 500 topics which are found to be the most important among the breakthrough patents. Though equation A.9 looks a bit like our construction of $\eta_{i,t}$ in equation 12, it differs in that we no longer directly use word vectors to compute similarities. Instead, the Cong et al. (2019) technique only uses the word vectors to give an educated guess on the topics contained in the set of documents. Thus occcupations are similar to a given topic when they contain words that are also found in that topic.

We focus on the period of time covered by our CPS merged outgoing rotation group sample (1985-2018) used in the employment regressions in Figure 8. This is for two reasons: first, this is the period where our employment and wage data coverage is most comprehensive, with a yearly time series and relatively stable occupation classifications. Second, the task composition of innovations has begun to change in this period of time relative to all previous innovation waves. In particular, cognitive skills have started to become more related to innovations, and this has been driven by the rising importance of information technology and electronics patents, which was not the case prior to the late 20th century. If skill-biased technological change has complemented the skillset of cognitive occupations, then innovations related to these occupations may be complementary to rather than a substitute for their skills. Thus if our measure mixes these two channels it is particularly likely to occur during this period of time.

Steps 1 and 2 above simply group documents into topics of related terms, compute how related a given topic is to each individual document, and provide an estimate of how important each topic is to the overall set of documents. Justification for the use of LSH clustering of word vectors to obtain topics and the singular value decomposition to infer topic importance/document topic loadings are discussed at length in Cong et al. (2019), to which we refer the interested reader for further details. For our purposes it suffices that by performing steps 1 through 3 we are able to obtain a panel of 500 predictors at the occupation-by-year level and which represent exposures to topics of words which are particularly relevant to patents.

In brief, this approach can be summarized as follows. We first extract the 500 most important common factors (topics) from the text of breakthrough patents using the approach of Cong et al. (2019) and the vector representations of word embedings discussed in Section 2. We then use these 500 textual factors to form a single predictor that is optimized to predict occupation declines in-sample. To do so, we examine the univariate performance of each factor in predicting employment declines, and then form a linear combination (the first principal component) of the predictors that are statistically significant negative predictors at the 5%. We also construct a labor-enhancing factor using the converse exercise.

Appendix Table A.7 summarizes our findings. By design, both factors predict employment with the correct sign in-sample. More importantly, both of these factors predict wage growth with the same sign, despite the fact that they were not designed to do so and wage growth is not highly correlated with employment growth. That said, the displacement factor (the factor calibrated for employment declines) is a much stronger predictor of both employment and wage growth than its counterpart designed to predict positive employment growth.

In terms of magnitudes, the employment and wage declines predicted by this statistical displacement factor are comparable to our baseline measure—that is, 1.25% vs 1.12% employment declines at the 10-year horizon and 0.25% vs 0.20% decline in wage earnings). The correlation between our baseline measure $\eta_{i,t}$ and the statistical predictor constructed to represent exposure to labor-saving technologies is approximately 73 percent. By contrast, the correlation with the factor calibrated to predict employment increases is negative at -11 percent.

A.7 Model Appendix

The model we use consider a continuum of workers, with a state parameter θ on the [0, 1] interval, corresponding to their ability to produce H. Share s_h of workers have the ability to accumulate H over time, and have H reset by a certain amount when new technology enters the playing field. Technology is given by ξ and has shocks of size κ, which (in expectation) displaces the human capital of α share of workers, reducing their θ by m.

Production is given by a nested CES production function, where composite good X is produced via a combination of L and ξ

$$X = \left(\xi^{\rho}\lambda + L^{\rho}(1-\lambda)\right)^{(1/\rho)}.$$

Output, Y, is produced as a combination of X and H:

$$Y = \left(\mu H^{\sigma} + (1-\mu)X^{\sigma}\right)^{(1/\sigma)}.$$

Technology evolves with process

$$\xi_t = (1-g)\xi_{t-1} + \kappa\, d\, N_t$$

where dN is a random variable with expectation ω.

Worker $i \in s_h$ has evolving human capital such that

$$\theta_{i,t} = m\,\theta_{i,t-1}\, dM_{i,t} - d_{i,t}\, h\, \theta_{i,t-1}.$$

In this case, dM is a random variable representing human capital acquisition, m is the size of the jump relative to initial human capital, dN is the same shock variable as in the equation for technology, and h_{-} is the scale of the loss of human capital (H) if a shock occurs. $d_{i,t}$ is an i.i.d. binomial random variable with expectation $E(d) = \alpha$, indicating whether someone is "exposed" to a technology shock or not. If you are exposed to a technology shock when one occurs, you experience the human capital loss, otherwise you do not.

We solve the model in discrete time, with a monthly time-step δt, and we approximate the continuum of workers with an exponentially increasing, finite grid of points on the [0, 1] interval. Since we are approximating a continuum of workers, each gridpoint has an infinite number of observations, and we can work directly with expectations when solving the model. This means for a given starting grid point on the θ interval, we have

$$E(\theta_{i,t}) = \theta_{i,t-1} + m\,\phi\,\theta_{i,t-1} - \alpha\, h\, \theta_{i,t-1}$$

Note that in discrete time, the technology and human capital processes admit a two-state Markov-Switching VAR representation with a shock state ($s = 0$) and a no-shock state $s = 1$. Let i index the starting gridpoint of a worker, with $i+1$ and $i-1$ being the adjacent gridpoints.

With some abuse of notation, instead of thinking of $\theta_{i,t}$ as a single worker on the grid, we can think of it as the probability mass (share of workers) on a gridpoint i.

$$E(\theta_{i,t}|s=0) = \theta_{i,t-1} - \phi\theta_{i+1,t-1} + \phi\theta_{i-1,t-1}$$

This works because we set the distance between gridpoints is m. If we have a shock, we have

$$E(\theta_{i,t}|s=1) = \theta_{i,t-1} - \phi\theta_{i+1,t-1} + \phi\theta_{i-1,t-1} - \alpha\theta_{i,t-1}) + \alpha\, h\, \theta_{i+m,t-1}).$$

In other words, if we experience a shock, we get some mass from the gridpoint which is mh above us, lose share α of our previous density as exposed workers, lose share ϕ to a higher gridpoint as workers acquire new skills, and gain share ϕ from the gridpoint below.

We can represent this as a VAR process, with transition probability α to a gridpoint which is

mh_- points below the current, ϕ to a gridpoint above us, and so on. In practice, in order to make h_- a continuous parameter, we split the fall probability across the two relevant gridpoints, with density allocated between them to make the fall have expectation h_-. For example, if m was 3%, and we needed a 5% fall, conditional on the fall a worker would have (roughly) a 2/3 chance of falling two gridpoints and a 1/3 chance of falling 1 gridpoint.

The transition process for the workers in s_l is very simple, as it is an absorbing state with no entry or exit. So

$$\theta_{s_l,t} = \theta_{s_l,t-1}$$

in both shock periods and no-shock periods.

The transition process for ξ is given by

$$\xi_t | s_t = 0 = (1-g)\xi_{t-1}$$

and

$$\xi_t | s_t = 1 = (1-g)\xi_{t-1} + \kappa.$$

Suppose we set up the VAR coefficient matrices, A, accordingly. Each period has probability ω of experiencing a shock, and probability $1-\omega$ of not experiencing a shock. This gives us transition process

$$E(A_t) = (1-\omega)A_{t-1,0} + \omega A_{t-1,1}.$$

Bianchi (2016) demonstrates how to find the steady state of the Markov-Sitching VAR model. For this exposition, we rely on his notation. He considers the MS-VAR process

$$Z_t = c_{\xi_t} + A_{\xi_t} Z_{t-1} + V_{\xi_t}\epsilon_t,$$

and

$$V_{\xi_t} = R_{\xi_t}\Sigma_{\xi_t},$$

where z_t is a vector of variables, c_t is a vector of constants,

In practice, our process for the mass of θ has a reflecting barrier at 1, and in order to enforce the sum-to-1 constraint for the total mass on the θ grid we represent the VAR through a VECM process.

We then solve for the steady state by finding the largest real eigenvalue of the Bianchi representation of the MS-VAR system. Because we have an absorbing state at the bottom of the θ grid whose density is a known value (fixed before calibration), we exclude that point from the solution, scaling down the intercept term by $1-s_l$. Finally, to compute the value at the top of the θ grid (call it gridpoint j), we simply compute $1-\sum_{i\ j}\theta_{ss,i} - s_l$.

Once we've solved for the ergodic steady state, we can begin calculating the model moments which correspond to our empirical calibration targets. Our process sets time step δt to one month.

Our theoretical moments are calculated as one-month impact responses, scaled by an annualization factor $\sqrt{12\omega(1-\omega)}$ for aggregate moments (labor share and output), and $\sqrt{12\omega\alpha(1-\omega\alpha)}$.

Impact responses for labor share and output are calculated relative to the ergodic steady state for all state variables. The steady state values are iterated forward one period using the transition matrices constructed above, but where a shock happens with certainty (effectively setting $\omega = 1$ for a single period). To compare output, labor share, wages, and other desired targets, between the values at the ergodic steady state relative to the shock period, we follow this procedure in each period. First, we calculate the level of H at the steady state as

$$H = \sum_{i=1}^{N} \theta_i m(\theta_i)$$

where $m(theta_i)$ is the mass of theta at gridpoint i. The workers in s_l produce no H, so this is sum of the value for θ at each gridpoints times the mass of workers at that rung of the ladder. L is calculated as $1 - H$. Output Y and the composite good X are calculated with the equations provided above. σ, ρ, μ, and λ are free parameters. If we call ξ^* the value of the technology state variable at the ergodic steady state, ξ post-shock is $\xi^* + \kappa$, where κ is a free variable. Wages associated with H (w_h) and L (w_l) are calculated as the marginal product of each task. Given output, these are calculated as

$$w_h = \frac{(1-\mu)\mu H^{\sigma-1}(\lambda\xi^\rho - L^\rho(\lambda-1))^{(\sigma/\rho)} + \mu H^\sigma)^{(1/\sigma)}}{(1-\mu)(\lambda\xi^\rho - L^\rho(\lambda-1))^{(\sigma/\rho)} + \mu H^\sigma},$$

and

$$w_l = \frac{(1-\mu)(\lambda\xi^\rho - L^\rho(\lambda-1))^{(\sigma/\rho)} + \mu H^\sigma)^{(1/\sigma)}(\lambda-1)(\lambda\xi^\rho - L^\rho(\lambda-1))^{(\sigma/\rho)}(\mu-1)L^{\rho-1}}{(\lambda\xi^\rho + (1-\lambda)L^\rho)((1-\mu)(\lambda\xi^\rho - L^\rho(\lambda-1))^{(\sigma/\rho)} + \mu H^\sigma)}$$

For each period in question for the impact responses, wages are calculated by plugging in the relevant state variables. Impact responses are calculated in log differences to align with our calibration targets (e.g. $\log Y_{shock} - \log Y_{ss}$), and subsequently scaled by the annualization factor.

Appendix Figures and Tables

Table A.1: Most Similar Patents For Select Occupations

Cashiers (SOC Code 412011)	
5055657	Vending type machine dispensing a redeemable credit voucher upon payment interrupt
5987439	Automated banking system for making change on a card or user account
5897625	Automated document cashing system
6012048	Automated banking system for dispensing money orders, wire transfer and bill payment
5598332	Cash register capable of temporary-closing operation
Loan Interviewers and Clerks (SOC Code 434131)	
6289319	Automatic business and financial transaction processing system
5611052	Lender direct credit evaluation and loan processing system
6233566	System, method and computer program product for online financial products trading
5940811	Closed loop financial transaction method and apparatus
5966700	Management system for risk sharing of mortgage pools
Railroad Conductors (SOC Code 534031)	
5828979	Automatic train control system and method
6250590	Mobile train steering
3944986	Vehicle movement control system for railroad terminals
6135396	System and method for automatic train operation
5797330	Mass transit system

Table A.2: Most Similar Occupations For Select Patents

“Knitting-machine” (Patent No. 276146, Issued in 1883)
Textile Knitting and Weaving Machine Setters, Operators, and Tenders
Sewing Machine Operators
Sewers, Hand
Fabric Menders, Except Garment
Textile Winding, Twisting, and Drawing Out Machine Setters, Operators, and Tenders
“Metal wheel for vehicles” (Patent No. 1405358, Issued in 1922)
Automotive Service Technicians and Mechanics
Cutting, Punching, and Press Machine Setters, Operators, and Tenders, Metal and Plastic
Maintenance Workers, Machinery
Grinding, Lapping, Polishing, and Buffing Machine Tool Setters, Operators, and Tenders, Metal and Plastic
Rolling Machine Setters, Operators, and Tenders, Metal and Plastic
“System for managing financial accounts by a priority allocation of funds among accounts” (Patent No. 5911135, Issued in 1999)
Financial Managers
Credit Analysts
Loan Interviewers and Clerks
Accountants and Auditors
Bookkeeping, Accounting, and Auditing Clerks

Table A.3: Occupations Most and Least Exposed to Innovation

Top 5 Occupations by Average $\eta_{i,t}$	Bottom 5 Occupations by Average $\eta_{i,t}$
Inspectors, Testers, Sorters, Samplers, and Weighers	Mental Health Counselors
Metal Workers and Plastic Workers, All Other	Dancers
Cutting, Punching, and Press Machine Setters, Operators, and Tenders, Metal and Plastic	Funeral Attendants
Production Workers, All Other	Judges, Magistrate Judges, and Magistrates
Electromechanical Equipment Assemblers	Clergy

Table A.4: Unconditional Correlations of $\eta_{i,t}$ With Task Categories

NR Cog (Analytical)	-0.12**					
	(-2.53)					
NR Cog (Interpersonal)		-0.16***				
		(-4.65)				
NR Man (Physical)			0.24***			
			(5.65)			
NR Man (Interpersonal)				-0.33***		
				(-8.43)		
Routine Cognitive					0.033	
					(0.95)	
Routine Manual						0.24***
						(5.74)

This figure plots the correlations of $\eta_{i,t}$ with the occupation task types computed from O*NET in Acemoglu and Autor (2011). Correlations are weighted by the Acemoglu and Autor (2011) occupation employment weights used to normalize the distribution of tasks to mean zero and standard deviation one.

Table A.5: Technology And Employment Over the Long Run (1850-2010)–Subsamples

	A. Early Sample				B. Later sample			
	Occ. 1850–1920		Occ.×Ind. 1910–1950		Occ. 1930–1990		Occ.×Ind. 1950–1990	
Technology Exposure, $\eta_{i,t}$	-0.66**	-0.60**	-1.10***	-1.41***	-0.80***	-0.69***	-0.59***	-0.64***
	(-2.19)	(-2.24)	(-3.61)	(-3.65)	(-4.13)	(-4.39)	(-3.00)	(-3.45)
Observations	968	776	25,712	14,945	1,606	1,432	55,297	39,717
R^2 (Within)	0.027	0.031	0.021	0.036	0.054	0.124	0.009	0.011
Controls								
Time FE	Y	Y			Y	Y		
Industry X Time FE			Y	Y			Y	Y
Lagged Dependent Variable		Y		Y		Y		Y

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

$$\frac{1}{k}\Big(\log Y_{i,t+k} - \log Y_{i,t}\Big) = \alpha_0 + \alpha_t + \beta(k)\eta_{i,t} + \rho\,(\log Y_{i,t} - \log Y_{i,t-k}) + \epsilon_{i,t}$$

for $k = 10, 20$ years for Census years spanning from 1850–2010. Here $Y_{i,t}$ is the occupation i's share in total non-farm employment in Census year t. The main variable of interest is $\eta_{i,t}$, our technology exposure measure (normalized to unit standard deviation). Employment growth rates is in annualized percentage terms. Standard errors are clustered by occupation and corresponding t-stats are shown in parentheses. Observations are weighted by occupation employment share at time t. Census year 1870 does not show up in the first column of the 20-year subsample regressions because the 1890 Census records no longer exist.

Table A.6: Technology And Employment During and Outside of Innovation Waves

	Innovation Wave	Other Years
Technology Exposure, $\eta_{i,t}$	-0.82***	-0.53
	(-5.92)	(-1.54)
Time FE	X	X
N	1106	1468
R^2 (Within)	0.091	0.018

The table above plots results from regressions of the form

$$\log(Y_{i,t+k}) - \log(Y_{i,t}) = \alpha_0 + \alpha_t + \beta\eta_{i,t} + \epsilon_{i,t}$$

for $k = 20$ years for Census years spanning from 1850-2000. Here $Y_{i,t}$ is occupation's share in total non-farm employment. $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. The sample is split into periods of innovation waves as identified by the breakthrough patent index of Kelly et al. (2020). The 20 year periods beginning in years 1880, 1910, 1920, and 1980, 1990 are labelled innovation waves with the remaining years representing non-innovation wave periods. Standard errors are clustered by occupation and corresponding t-stats are shown in parentheses. Observations are weighted by occupation employment share at time t.

Figure A.1: Sample Patent Topic Word Clusters

The above are four of the topics resulting from the LSH approximate nearest neighbors routine used to separate words into clusters as described in section 3.3. The relative size of the word corresponds to the importance of that word within the topic.

Table A.7: Predictive Performance of 10-Year Employment and Wage Growth on Predictors Constructed From Patent Topics

Panel A: Negative Constructed Predictor

	Employment Growth		**Wage Growth**	
ξ_{Mean}	-1.25***		-0.21***	
	(-6.79)		(-6.10)	
ξ_{PC1}		-1.09***		-0.20***
		(-6.32)		(-6.11)
Year FEs	X	X	X	X
Controls	X	X	X	X

Panel B: Positive Constructed Predictor

	Employment Growth		**Wage Growth**	
γ_{Mean}	0.59***		0.019	
	(3.84)		(0.70)	
γ_{PC1}		0.50***		0.0062
		(3.67)		(0.24)
Year FEs	X	X	X	X
Controls	X	X	X	X

The tables above show coefficients from panel regressions of annualized wage and income growth rates over the 10-year horizon on textual factors constructed to predict employment as described in section 3.3. Regressions are of the form

$$y_{i,t+k} - y_{i,t} = \alpha + \beta Z_{i,t} + \delta X_{i,t} + \varepsilon_{i,t}$$

For $Z_{i,t} = \xi_{i,t}$ ("labor-saving") or $\gamma_{i,t}$ ("productivity enhancing"). Controls $X_{i,t}$ include three one-year lags of dependent variable, time fixed effects, wage, and occupation employment share. Subscripts $PC1$ and $Mean$ denote versions computed using either the first principal component or cross-sectional mean across individual textual predictors derived from the patent topics identified by the Cong et al. (2019) method. Dependent variable is expressed in annualized percentage terms and $\eta_{i,t}$ is standardized. Standard errors are clustered by occupation and independent variables are standardized. Observations are weighted by occupation's employment share at time t. The sample uses CPS merged outgoing rotation group data starting in 1982.

Table A.8: Correlations Between Predictors Constructed From Patent Topics and Different Versions of Occupation Technology Exposure $\eta_{i,t}$

	All Patents		Drop ICT Patents		Just ICT Patents	
ξ_{Mean}	0.73***		0.88***		0.41***	
	(20.53)		(31.71)		(10.21)	
γ_{Mean}		-0.11***		-0.17***		-0.030
		(-5.26)		(-6.39)		(-1.13)

This table reports correlations between versions of technology exposure $\eta_{i,t}$ formed using different sets of patents and the composite predictors constructed from textual factors using the Cong et al. (2019) method to predict employment outcomes either negatively (ξ_{Mean}) or positively (γ_{Mean}). The "$Mean$" label denotes versions of composite predictors constructed by taking the cross-sectional means across individual textual factors which predict employment either negatively or positively. The first two columns represent the baseline measure of $\eta_{i,t}$ constructed using all patents; the next two columns drop ICT patents, defined to be those falling under the instruments/information or electronics categories; finally, the last two columns form $\eta_{i,t}$ only using ICT patents.

Table A.9: Breakthrough patents most related to tasks performed by order-fulfillment clerks

US. Pat. #	Distance ($\tilde{\rho}$)	Issue Year	Title
5,696,906	0.933	1997	Telecommunication user account management system and method
5,627,973	0.915	1997	Method and apparatus for facilitating evaluation of business opportunities for supplying goods and/or services to potential customers
5,689,705	0.896	1997	System for facilitating home construction and sales
5,592,560	0.885	1997	Method and system for building a database and performing marketing based upon prior shopping history
5,687,212	0.885	1997	System for reactively maintaining telephone network facilities in a public switched telephone network
5,628,004	0.881	1997	System for managing database of communication of recipients
5,621,812	0.880	1997	Method and system for building a database for use with selective incentive marketing in response to customer shopping histories
5,638,457	0.880	1997	Method and system for building a database for use with selective incentive marketing in response to customer shopping histories
5,659,469	0.879	1997	Check transaction processing, database building and marketing method and system utilizing automatic check reading
5,592,378	0.874	1997	Computerized order entry system and method
5,787,405	0.896	1998	Method and system for creating financial instruments at a plurality of remote locations which are controlled by a central office
5,802,513	0.884	1998	Method and system for distance determination and use of the distance determination
5,717,596	0.878	1998	Method and system for franking, accounting, and billing of mail services
5,797,002	0.873	1998	Two-way wireless system for financial industry transactions
5,812,985	0.866	1998	Space management system
5,774,877	0.866	1998	Two-way wireless system for financial industry transactions
5,848,396	0.865	1998	Method and apparatus for determining behavioral profile of a computer user
5,790,634	0.865	1998	Combination system for proactively and reactively maintaining telephone network facilities in a public switched telephone system
5,734,823	0.864	1998	Systems and apparatus for electronic communication and storage of information
5,712,987	0.864	1998	Interface and associated bank customer database
5,995,976	0.912	1999	Method and apparatus for distributing supplemental information related to printed articles
6,006,251	0.897	1999	Service providing system for providing services suitable to an end user request based on characteristics of a request, attributes of a service and operating conditions of a processor
5,930,764	0.889	1999	Sales and marketing support system using a customer information database
5,884,280	0.886	1999	System for and method of distributing proceeds from contents
5,991,728	0.884	1999	Method and system for the tracking and profiling of supply usage in a health care environment
5,903,873	0.876	1999	System for registering insurance transactions and communicating with a home office
5,991,876	0.875	1999	Electronic rights management and authorization system
5,953,389	0.869	1999	Combination system for provisioning and maintaining telephone network facilities in a public switched telephone network
5,893,075	0.868	1999	Interactive system and method for surveying and targeting customers
5,932,869	0.867	1999	Promotional system with magnetic stripe and visual thermo-reversible print surfaced medium
6,041,319	0.876	2000	Method and system for telephone updates of postal scales
6,061,506	0.874	2000	Adaptive strategy-based system
6,072,493	0.869	2000	System and method for associating services information with selected elements of an organization
6,105,003	0.864	2000	Customer data processing system provided in a showroom
6,070,160	0.854	2000	Non-linear database set searching apparatus and method
6,023,705	0.854	2000	Multiple CD index and loading system and method
6,154,753	0.846	2000	Document management system and method for business quality modeling
6,064,879	0.845	2000	Mobile communication method, and mobile telephone switching station customer management system, and mobile unit for implementing the same
6,112,181	0.842	2000	Systems and methods for matching, selecting, narrowcasting, and/or classifying based on rights management and/or other information

Table A.10: Worker Earnings and Technology Exposure, by Prior Income and Education / Age

	(1)	(2)	(3)	(4)	(5)
Income Percentile		Age		Education	
	25–35	35–45	45–55	College	No Coll
A. Cond. Mean, $E[g]$, by Horizon					
0 to 25-th	-1.05	-1.72	-2.03	-2.14	-1.11
	(-2.57)	(-6.36)	(-5.06)	(-6.32)	(-3.58)
25 to 50-th	-0.28	-0.87	-2.24	-0.92	-1.11
	(-0.97)	(-2.89)	(-3.18)	(-2.72)	(-4.40)
50 to 75-th	-0.57	-0.68	-2.43	-0.82	-1.60
	(-1.46)	(-2.09)	(-3.32)	(-2.02)	(-4.36)
75 to 95-th	-0.39	-0.49	-2.22	-0.60	-2.13
	(-0.53)	(-1.56)	(-4.73)	(-1.37)	(-7.52)
95 to 100-th	-3.07	-1.85	-2.86	-1.98	-3.50
	(-4.22)	(-3.07)	(-3.51)	(-4.10)	(-4.47)
B. Risk, Absolute Income Growth $E[\|g\|]$					
0 to 25-th	0.01	-0.55	0.06	-0.51	-0.18
	(0.04)	(-2.13)	(0.18)	(-1.74)	(-0.77)
25 to 50-th	0.20	-0.17	0.55	0.12	0.03
	(0.68)	(-0.81)	(0.83)	(0.48)	(0.11)
50 to 75-th	0.42	0.06	1.17	0.50	0.49
	(1.66)	(0.31)	(1.86)	(1.85)	(1.44)
75 to 95-th	0.61	0.12	1.51	0.66	1.01
	(1.56)	(0.45)	(3.84)	(2.87)	(2.94)
95 to 100-th	2.38	1.44	2.38	2.01	2.09
	(3.05)	(3.03)	(3.92)	(6.02)	(2.76)
C. Skewness, Prob. Large Income Decline $p(g < p^{10})$					
0 to 25-th	0.18	0.25	0.68	0.40	0.20
	(0.99)	(1.93)	(4.37)	(2.5)	(1.39)
25 to 50-th	0.07	0.16	0.73	0.24	0.28
	(0.3)	(0.86)	(2.42)	(1.37)	(1.34)
50 to 75-th	0.18	0.16	0.91	0.26	0.56
	(1.26)	(1.08)	(3.05)	(1.82)	(4.03)
75 to 95-th	0.07	0.05	0.90	0.15	0.83
	(0.28)	(0.27)	(3.99)	(0.87)	(6.28)
95 to 100-th	2.42	0.68	1.46	1.14	1.53
	(4.52)	(2.24)	(4.43)	(4.56)	(4.89)

Note: Panel A shows the estimated slope coefficients (times 100) from equation (21) in the main text: the dependent variable is worker earnings growth over a 5-year horizon; the main independent variable of interest is a worker's technology exposure $\eta_{i,t}$. The slope estimate $\beta(h)$ is allowed to vary with the worker's prior income rank and age (columns (1) to (3)) or education (columns (4) to (5)). Panel B shows the slope coefficients of a variant of the above specification where we replace the dependent variable $g_{i,t:t+h}$ with its absolute value $|g_{i,t:t+h}|$. Panel C replaces the dependent variable with a dummy that takes the value of one if $g_{i,t:t+h}$ lies in the bottom 10-th percentile. See notes to Tables 3 to 5 for additional details.

Table A.11: Worker Earnings and Technology Exposure, by Education

Horizon	Education	(1)	(2)	(3)	(4)
3 Years	College	-0.975	-0.880	-0.924	-0.911
		(0.26)	(0.246)	(0.292)	(0.274)
	No College	-1.32	-1.27	-1.30	-1.30
		(0.153)	(0.168)	(0.253)	(0.233)
	No College - College	-0.347	-0.391	-0.378	-0.388
		(0.209)	(0.197)	(0.199)	(0.192)
5 Years	College	-1.35	-1.17	-1.20	-1.14
		(0.291)	(0.282)	(0.362)	(0.343)
	No College	-1.68	-1.55	-1.54	-1.50
		(0.216)	(0.225)	(0.31)	(0.291)
	No College - College	-0.336	-0.387	-0.332	-0.3603
		(0.2315)	(0.225)	(0.230)	(0.225)
10 Years	College	-1.24	-1.08	-1.31	-1.21
		(0.352)	(0.343)	(0.504)	(0.486)
	No College	-1.82	-1.67	-1.85	-1.76
		(0.281)	(0.262)	(0.404)	(0.393)
	No College - College	-0.581	-0.581	-0.544	-0.550
		(0.283)	(0.274)	(0.290)	(0.283)
Fixed Effects:					
Industry		X	X		
Occupation		X		X	
Year		X			
Industry x Year				X	X
Occupation x Year			X		X

Table A.12: Worker Earnings and Technology Exposure, by Age

Worker Age	Horizon	(1)	(2)	(3)	(4)
3 Years	25-35	-0.479	-0.396	-0.389	-0.388
		(0.268)	(0.244)	(0.224)	(0.209)
	35-45	-0.903	-0.83	-0.866	-0.857
		(0.18)	(0.159)	(0.181)	(0.164)
	45-55	-1.98	-1.91	-1.96	-1.95
		(0.392)	(0.42)	(0.542)	(0.524)
	(45-55) - (25-35)	-1.50	-1.51	-1.57	-1.56
		(0.428)	(0.441)	(0.463)	(0.462)
5 Years	25-35	-0.874	-0.706	-0.683	-0.64
		(0.243)	(0.227)	(0.271)	(0.251)
	35-45	-1.23	-1.08	-1.09	-1.04
		(0.148)	(0.148)	(0.221)	(0.204)
	45-55	-2.44	-2.29	-2.3	-2.25
		(0.535)	(0.542)	(0.656)	(0.634)
	(45-55) - (25-35)	-1.56	-1.59	-1.62	-1.61
		(0.489)	(0.494)	(0.516)	(0.512)
10 Years	25-35	-0.986	-0.823	-1.01	-0.918
		(0.291)	(0.266)	(0.397)	(0.38)
	35-45	-1.43	-1.26	-1.47	-1.37
		(0.238)	(0.228)	(0.39)	(0.377)
	45-55	-2.54	-2.40	-2.60	-2.51
		(0.682)	(0.676)	(0.861)	(0.832)
	(45-55) - (25-35)	-1.56	-1.58	-1.59	-1.59
		(0.565)	(0.564)	(0.579)	(0.574)
Fixed Effects:					
Industry		X	X		
Occupation		X		X	
Year		X			
Industry x Year				X	X
Occupation x Year			X		X

Table A.13: Worker Earnings and Technology Exposure, by Prior Income

Horizon	Income Rank	(1)	(2)	(3)	(4)
3 Years	$[0, 25)$	-1.34	-1.21	-1.28	-1.24
		(0.236)	(0.239)	(0.207)	(0.216)
	$[25, 50)$	-0.934	-0.811	-0.883	-0.849
		(0.145)	(0.167)	(0.212)	(0.205)
	$[50, 75)$	-1.10	-0.974	-1.05	-1.01
		(0.241)	(0.244)	(0.304)	(0.289)
	$[75, 95)$	-1.14	-1.00	-1.09	-1.05
		(0.28)	(0.266)	(0.26)	(0.256)
	$[95, 100]$	-2.31	-2.20	-2.25	-2.24
		(0.421)	(0.422)	(0.374)	(0.366)
5 Years	$[0, 25)$	-1.76	-1.51	-1.56	-1.49
		(0.234)	(0.258)	(0.267)	(0.281)
	$[25, 50)$	-1.26	-1.02	-1.08	-1.01
		(0.186)	(0.21)	(0.281)	(0.282)
	$[50, 75)$	-1.43	-1.19	-1.24	-1.18
		(0.33)	(0.323)	(0.395)	(0.383)
	$[75, 95)$	-1.42	-1.17	-1.24	-1.17
		(0.328)	(0.315)	(0.345)	(0.339)
	$[95, 100]$	-2.71	-2.48	-2.52	-2.47
		(0.578)	(0.574)	(0.516)	(0.52)
10 Years	$[0, 25)$	-2.05	-1.74	-2.00	-1.85
		(0.36)	(0.337)	(0.395)	(0.393)
	$[25, 50)$	-1.14	-0.858	-1.1	-0.964
		(0.338)	(0.349)	(0.474)	(0.472)
	$[50, 75)$	-1.19	-0.90	-1.16	-1.01
		(0.414)	(0.40)	(0.578)	(0.541)
	$[75, 95)$	-0.986	-0.686	-0.967	-0.807
		(0.445)	(0.436)	(0.527)	(0.492)
	$[95, 100]$	-2.46	-2.20	-2.42	-2.28
		(0.589)	(0.594)	(0.611)	(0.596)
Fixed Effects:					
Industry		X	X		
Occupation		X		X	
Year		X			
Industry x Year				X	X
Occupation x Year			X		X

Table A.14: Technology And Employment Over the Long Run (1850-2010)–Heterogenous effects by age

	A. Full Sample	B. Sub-samples		
		1850–1920	1930–1960	1970–1990
Age (20–29) × Technology Exposure, $\eta_{i,t}$	-0.71***	-2.24***	-0.073	-0.58**
	(-3.46)	(-4.08)	(-0.19)	(-2.33)
Age (30–39) × Technology Exposure, $\eta_{i,t}$	-0.57***	-1.70***	-0.098	-0.50**
	(-3.41)	(-3.28)	(-0.31)	(-2.44)
Age (40–49) × Technology Exposure, $\eta_{i,t}$	-1.10***	-2.13***	-0.62*	-1.03***
	(-6.20)	(-4.40)	(-1.88)	(-4.85)
Observations	6,512	2,232	1,989	2,291
R^2 (Within)	0.066	0.074	0.055	0.090
Controls				
Age Group X Year FE	Y	Y	Y	Y
Lagged Dependent Variable	Y	Y	Y	Y
P-val (40–49) - (20–29)	0.015	0.776	0.129	0.002

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

$$\frac{1}{k}\Big(\log Y_{i,a',t+k} - \log Y_{i,a,t}\Big) = \alpha_0 + \alpha_t + \beta(k,a)\eta_{i,t} + \rho\left(\log Y_{i,t} - \log Y_{i,t-k}\right) + \epsilon_{i,t}$$

for $k = 20$ years for Census years spanning from 1850-2010, where we allow the coefficients to vary by age. The dependent variable $Y_{i,a,t}$ tracks employment by workers in age group a; for example, to compute the 20-year growth rate in employment in 1900 for workers aged 20–29 in occupation i, we compare it to the employment of 40–49 year old workers in occupation i in 1920. The main variable of interest is $\eta_{i,t}$, our technology exposure measure (normalized to unit standard deviation). Employment growth rates is in annualized percentage terms. Standard errors are clustered by occupation and corresponding t-stats are shown in parentheses. Observations are weighted by occupation employment share at time t. Census year 1870 does not show up in the first column of the 20-year subsample regressions because the 1890 Census records no longer exist.