Technology, Vintage-Specific Human Capital, and Labor Displacement: Evidence from Linking 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
Kellogg School of Management

First Version: December 2019
Current Version: July 2022

Abstract

We develop a granular, occupation-specific measure of technological progress that relies only on textual descriptions of patent documents and the tasks performed by workers in an occupation. Our measure primarily identifies labor-saving innovations and is broadly available from the 19th century to the present. 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. Using a panel of administrative data on worker earnings, we show that the earnings of older and more highly-paid workers are more responsive to our technology exposure measure, a pattern consistent with skill displacement. Our calibrated model fits these facts and emphasizes the importance of movements in skill quantities, not just skill prices, for the link between technology and inequality.

*We are grateful to Daron Acemoglu, Andy Atkeson, David Autor, Tania Babina, Effi Benmelech, Julieta Caunedo, Martin Beraja, Andrea Eisfeldt, Tarek Hassan, Michael Peters, Pascual Restrepo, Jonathan Rothbaum, Miao Ben Zhang, and seminar participants at University of Amsterdam, Boston University, FIRS, Josh Hopkins, HKUST, Labor and Finance Group, NBER EFG, MIT Sloan, Michigan State, the Econometric Society, Rice University, University of Rochester, the Society of Economic Dynamics, University College London, University of Illinois at Urbana Champaign, University of Toronto, Tsinghua PBC, WFA, and Wharton for valuable discussions and feedback. We thank Carter Braxton, Will Cong, and Jonathan Rothbaum for generously sharing code. Brice Green and Jinpu Yang provided excellent research support. The paper has been previously circulated under the titles "Technological Change and Occupations over the Long Run" and "Technology-Skill Complementarity and Labor Displacement: Evidence from Linking Two Centuries of Patents with Occupations." The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product (Data Management System (DMS) number: P-7503840, Disclosure Review Board (DRB) approval numbers: CBDRB-FY21-POP001-0176, CBDRB-FY22-SEHSD003-006, CBDRB-FY22-SEHSD003-023, CBDRB-FY22-SEHSD003-028).

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 as measurement lags theory.[3]

Our goal is to fill this gap: we leverage over a century 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 identify 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 (2021). To estimate the similarity 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.

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. To this end, we employ public-use Census micro-data that are available over longer horizons. When we compare workers across occupations differentially exposed to technology improvements, we find that an acceleration in the rate of breakthrough innovations is associated with future declines in employment and average wage earning. The negative correlation between employment and our technology exposure measure is largely consistent over time—starting

[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); Caunedo, Jaume, and Keller (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).

from the beginning of the 20th century to the present. This negative correlation between our technology exposure and both employment and wages suggests that out approach primarily identifies labor-saving innovations. To reinforce this view, we exploit recent advances in topic modeling to construct a composite predictor from the text of patents and occupations whose purpose is to maximize the in-sample predictability of employment declines–i.e., to identify language consistent with labor saving innovations. This statistical predictor likely represents an upper bound to how well one can predict employment declines using text from patents and occupations. Comparing the performance of our measure to this benchmark, we find that our measure predicts quantitatively similar declines in employment and earning.

We conclude that our indices capture the extent to which workers in a given occupation are exposed to breakthrough labor-saving innovations that arrive at a point in time. 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 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.

Armed with a measure of technology exposure, we next examine how it relates to subsequent worker earnings. We use confidential administrative earnings records from the US Social Security Administration, starting in the early 1990s, 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

[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."

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.

We find that, in response to a standard deviation increase in technology, the average worker experiences approximately a 0.02 log point decline in her wage earnings over the next five years. The point estimates are quite similar regardless of whether we control for common shocks to labor demand at the industry and occupation levels via industry-time and occupation-time fixed effects, respectively. Further, this relation is pervasive across broad sectors—magnitudes are similar across services and manufacturing. It is stronger in occupations that emphasize routine tasks and is weakest in occupations that emphasize interpersonal tasks. Further it is pervasive across education levels: we find no meaningful difference in earnings responses across workers with or without a four-year college degree.

Importantly, we find significant heterogeneity in these responses across age and income levels. In particular, older workers are significantly more affected than younger workers. In addition, workers at the top end of the earnings distribution—relative to their peers in the same occupation and in the same industry—experience a significantly greater decline in earnings (more than twice) relative to the average worker. One interpretation of this pattern is that firms selectively adopt labor-saving technologies to replace 'excessively' paid workers. However, these income pattern persists when we control for worker age, when we rank workers based on their income relative to the firm's average wage, or when we rank workers based on their income net of a battery of observables, including the location of their workplace (commuting zone) or unionization status. Further, we find no meaningful difference in this income gradient when we compare workers who are members of a labor union to those that are not. These patterns suggest that selective adoption is not the main driver of this fact.

Our interpretation of these pattern is that it is consistent with the importance of vintage-specific human capital. That is, improvements in technologies are often associated with obsolescence of certain worker skills—or simply with a period of learning during which productivity is depressed. To the extent that workers that are more highly paid relative to their peers have acquired more skills, we would expect such workers to experience lower wage growth as their skills become obsolete. Consistent with this view, we find that the differential response of high-paid workers relative to the lower-paid group is significantly stronger in occupations that require a greater amount of related experience.

At first, these facts may appear to be at odds with a common view that recent technological advances are skill-biased. Indeed, a common view is that the higher-paid workers perform tasks that are complementary to technology, so one might expect these workers to be experiencing earnings increases when related technologies emerge. Indeed, we show that this is the case when we focus on breakthrough technologies in the workers industry that have very low levels of similarity to the worker's tasks. Earnings of all workers subsequently increase, with workers at the top experiencing larger increases than the average worker. Thus, not all breakthrough technologies are labor-displacive; it depends on their similarity with the tasks that workers perform.

We rationalize our findings in the context of a model that features skill-biased technical change (Krusell et al., 2000) and vintage-specific human capital (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). In particular, 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 the facts in the data. 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 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. Interestingly, in both cases the behavior of the skill premium is an insufficient summary statistic for income inequality; technology moves both skill prices but also the quantity of skills.

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 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; Caunedo et al., 2021; 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.[5] 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.

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 and industries. 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.[6]

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

[6]See also Graetz and Michaels (2018); Dauth, Findeisen, Suedekum, and Woessner (2021); Koch, Manuylov, and

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 Measuring Workers' Technology Exposure

We begin our analysis by constructing a measure of workers' exposure to important technological innovations at a given point in time. There are two ingredients in this construction. The first part is the definition of what constitutes an important technological innovation. We rely on patent data and follow the methodology of Kelly et al. (2021), henceforth KPST, to identify important innovations. KPST identify breakthrough innovations as those that are both novel (whose descriptions are distinct from their predecessors) and impactful (they are similar to subsequent innovations). In particular, KPST first create a measure of importance for each patent that combines novelty and impact and then define a 'breakthrough' patent as one that falls in the top 10% of the distribution of importance. KPST show that these breakthrough technologies are associated with increases in measured productivity both at the aggregate as well as the industry level.

The second part involves identifying the set of workers who are most exposed to a particular technological breakthrough. To do so we rely on the description of job tasks a given occupation performs; for each breakthrough innovation we then construct a distance metric 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, or DOT). This section briefly describes this process; see Appendix A.2 for further details.

1.1 Data

Our goal is to identify breakthrough technologies that are related to the tasks performed by specific workers. We do so by analyzing the textual similarity between the description of the innovation in

Smolka (2021); Aghion, Antonin, Bunel, and Jaravel (2021); and Bessen, Goos, Salomons, and van den Berge (2022) for more recent work on the causes and effects of adopting robots and other automation technologies.

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 from 1991.[7] Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we use a crosswalk from DOT occupations to the considerably coarser and yet still detailed set of 6-digit Standard Occupation Classification (SOC) codes from O*NET.[8] We then combine all tasks for a given occupation at the 2010 SOC 6-digit level into one occupation-level corpus. We use the patent text data from Kelly et al. (2021) and combine the claims, abstract, and description section into one patent-level corpus for each patent.

1.2 Similarity between occupation tasks and patents

The language and terminology used in patent documents is quite different from occupational task descriptions. Consequently, standard methods which require exact overlap in terms (such as the "bag-of-words" approach described in Gentzkow, Kelly, and Taddy (2019)) are likely to perform relatively poorly in measuring patent–occupation textual similarity. Instead, we leverage recent advances in natural language processing that do not require exact overlap in terminology, and hence allow for synonyms or other relatedness in word meanings. Specifically, we use word embeddings—which capture geometric representations of word meanings in the form of dense vectors—as the basis for our occupation–patent textual comparisons. Word embeddings (also called word vectors) are estimated such that the distance between two word vectors is directly 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. Appendix Section A.2 contains further details for the procedure outlined here, along with a brief discussion of how the Pennington et al. (2014) word embeddings are constructed.

We represent each document A_i (either a patent or an occupation description) as a vector X_i, constructed as a weighted average of the set of word vectors x_k for the terms contained in the document:

$$X_i = \sum_{x_k \in A_i} w_{i,k} x_k. \tag{1}$$

A key part of the procedure consists of choosing appropriate weights $w_{i,k}$ in order to emphasize important words in the document. We follow the literature on natural language processing and

[7]Given that much of the empirical analysis in this paper focuses on worker-level outcomes in the post-1980 sample, we use the task descriptions from the earlier 1991 DOT, rather than task descriptions from O*NET that were released much more recently.

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

construct weights based on 'term-frequency-inverse-document-frequency' (TF-IDF). In brief, TF-IDF 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.

Last, we next use the cosine similarity to measure the similarity between patent i and occupation j,

$$\text{Sim}_{i,j} = \frac{X_i}{||X_i||} \cdot \frac{X_j}{||X_j||} \tag{2}$$

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 job tasks.[9]

Our analysis so far delivers a measure of similarity (2) between a given patent j and a given occupation i. We perform two adjustments to this measure. First, we remove yearly fixed effects. We do so in order to account for language and structural differences in patent documents over time.[10] Second, we impose sparsity: after removing the fixed effects we set all patent $\times$ occupation pairs to zero that are below the 80th percentile in this fixed-effect adjusted similarity. This imposes that the vast majority of patent–occupation pairs are considered unrelated to one another, and only similarity scores sufficiently high in the distribution receive any weight. Last, we scale the remaining non-zero pairs such that a patent/occupation pair at the 80th percentile of yearly adjusted similarities has a score equal to zero and the maximum adjusted score equals one. We denote by $\rho_{i,j}$ the adjusted similarity metric between patent j and occupation i. While we compute $\rho_{i,j}$ in this manner for all patent–occupation pairs, in the analysis that follows we restrict to the set of patents identified as breakthroughs by the KPST procedure.

[9] 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 1840.

[10] 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.

2 Validation and Interpretation

The next step in our analysis is to verify that our measure indeed captures exposure of workers to related technologies. We do so in two ways. First, we examine specific examples of breakthrough technologies and identify the most related occupations. Second, we explore the ability of our technology exposure measure to predict changes in employment and wages at the occupation level. We find that our measure strongly predicts employment and wage declines.

In addition, we explore the extent to which our exposure measure mostly captures labor-saving technologies. Given that our technology exposure measure is constructed based on the similarity between patents and tasks performed by a specific occupation, it is not entirely obvious ex-ante 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 find a consistently negative relation between our technology exposure measure and subsequent labor market outcomes, it is possible these effects are muted because our measure mixes labor-saving and labor-enhancing innovations; however, in this section we present evidence that such counteracting effects are likely to be small, and conclude that our measure primarily captures labor-saving innovations as intended.

2.1 Examples

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. To illustrate the effectiveness of our methodology in identifying links between technology and occupation task descriptions, we consider a three representative examples of breakthrough patents in Figure 1. 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 breakthrough 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 "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. Many of these technologies are labor-saving. 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.

Most broadly, 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. (2021). 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.[11] Both of these patents are in the list of breakthrough patents identified by Kelly et al. (2021). 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 highly skilled workers.

Last, we can also examine more recent examples of labor-saving wages and directly relate them to wage earnings. 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 1997 to 2002 as a period featuring a significant uptick in 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.5 contains a longer list of some of the most related breakthrough patents to order fulfillment clerks issued in the 1997 to 2000 period. Appendix Figure A.2 shows that, relative to all other clerk occupations, wage trends were fairly flat prior to 1997. However, since then they begin to systematically diverge following the arrival of significant breakthrough innovations that were related more to order clerks than other clerk occupations. By 2010, the log wages of order clerks had declined by about 20% (0.2 log points), relative to the difference in average log wages between the two groups in 1997.

[11] Jerome (1934) summarizes their impact: "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."

2.2 Relation with employment and wage growth

We next examine the relation between our technology exposure measure and subsequent growth in the employment shares and average wage earnings of exposed occupations. To this end, we construct a time series index of exposure of occupation i[12] to technology at time t as

$$\eta_{i,t} = \frac{1}{N_t} \sum_{j \in \mathcal{B}_t} \rho_{i,j}. \tag{3}$$

In words, $\eta_{i,t}$ aggregates our patent-occupation similarity scores $\rho_{i,j}$ across all breakthrough patents issued in period t. It varies over time due to the arrival of breakthrough technologies and it varies across occupations as these breakthrough technologies have different levels of similarity with the tasks performed by each occupation. We scale this measure by US population N_t so that it is stationary over the long period of time in consideration. The KPST breakthrough patents series ends in 2002, and the earliest Census year for which we can obtain occupation information is 1850, so the sample period for $\eta_{i,t}$ covers 1850-2002.

Technology exposure and employment growth

We begin by examining the extent to which our time-series exposure measure (3) is related to occupation employment using the Decennial Census. The Census surveys consist of repeated cross-sectional observations. Important for our purposes they contain information on occupations, which we can link to our technology exposure measure. The data consists of an unbalanced panel of occupation–Census year employment shares and spans the Census decades from 1910 to 2010.[13] Appendix A.4 provides details on data construction.

We examine employment outcomes using the following specification,

$$\frac{100}{h}\Big(\log Y_{i,t+k} - \log Y_{i,t}\Big) = \alpha_0 + \beta(h)\,\eta_{i,t} + \delta Z_{t,j} + \varepsilon_{i,t}, \qquad h = 10, 20 \text{ years}. \tag{4}$$

The variable $Y_{i,t}$ is the employment share in total non-farm employment of occupation i, and so the dependent variable is the annualized percentage growth rate in employment. Observations are weighted by the employment share of the given occupation and standard errors are clustered by occupation. Given that the time period t refers to a decade, we use the average value of $\eta_{i,t}$ over the preceding ten years for each period t and we normalize it to unit standard deviation. All

[12] While we originally compute textual similarities using DOT documents defined at the O*NET SOC code level, in the remainder of the paper we crosswalk SOC occupations to the revised Census occ1990 level (commonly referred to as “occ1990dd" codes) from Autor and Dorn (2013). Accordingly, we now compute $\rho_{i,j}$ as the average adjusted similarity score across all SOC codes (as calculated in section 1.2) that are linked to an individual occ1990dd occupation code.

[13] We start in 1910 because this was the first Census year to offer separate entries for activities now designated as industry and occupation. See https://usa.ipums.org/usa/volii/91occtc.shtml.

specifications include time fixed effects; depending on the specification, we include controls for the lagged 10-year employment growth rate.

Table 1 summarizes our findings. Examining Panel A, we note 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 0.55 to 0.64 percentage point annualized decline annualized decline in employment over the next 10 years. Extending the horizon over the next 20 years increases the magnitudes to a 0.77 to 0.87 percentage point annualized decline which corresponds to a cumulative decline in occupation employment of approximately 15 to 17 percent.

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 and estimate a modified version of equation (4), where now the dependent variable 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, we see that controlling for industry-specific time trends leads to quantitatively similar results. 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.

Technology exposure and wage growth

One potential concern with the previous analysis is that what we are picking up is the endogeneity of technological change: innovative effort is directed towards occupations for which labor supply is shrinking. However, if that were the case, we would expect to see that our technology measure would pick up wage increases (since it responds to a negative labor supply shock).

To examine the response of wages, we rely on 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 create a balanced panel of wage earnings and employment growth at the level of occupation and calendar year. We estimate specification similar to(4), over horizons k of 5 to 20 years. The dependent variable $Y_{i,t}$ now represents the average wage earnings or total 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 unit standard deviation.

Figure 4 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. 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; this estimate is quantitatively similar to the estimated decline in employment reported in Table 1.

More importantly, our innovation measure predicts a significant decline in occupation-level average wage earnings: a one-standard deviation increase in $\eta_{i,t}$ is followed by a decline in average wage earnings of approximately 0.2% per year. The fact that average wages at the occupation level decline following increases in our technology exposure measure is inconsistent with the view that the negative correlation with employment growth is driven by directed technical change in occupations that (are predicted to) experience labor scarcity.

2.3 Comparison to a purely statistical predictor

The results in the previous section illustrate that our technology exposure is strongly predictive of employment and wage declines at the occupation level. As such, it appears that our technology exposure measure is primarily identifying labor-saving technologies. Ex-ante, this conclusion is not obvious: the similarity between the description of an innovation and occupation tasks could in principle also capture technologies that complement the productivity of incumbent workers. This raises the question of whether the effects we identify are muted because our measure mixing labor-saving and labor-enhancing innovations.

One way to answer this question is to construct a pure machine-learning predictor of employment declines using the text of patents and compare its in-sample performance to the performance of our exposure measure. We construct this in-sample predictor by leveraging recent advances in topic modeling, building on the approach proposed by Cong, Liang, and Zhang (2019), which we briefly describe below. For more details on the procedure refer to Appendix A.6. To keep the exercise simple, we estimate the set of topics for patents issued in the more recent post-1976 period for which the USPTO provides full patent textual information.

We use the approach of Cong et al. (2019) to generate thousands of potential textual factors (topics), and then we extract the 500 most important topics from the patent texts. To do this we compute document topic "loadings" (a measure of how often the topic shows up in a document). While we compute topic loadings for both all breakthrough patents and all occupations, we only use the patent text to determine which 500 candidate topics to select. We then create year-t occupational exposures to each topic by multiplying the occupational loading on the kth topic by the sum of year-t breakthrough patent loadings on topic k. This yields 500 candidate textual predictors of employment declines. Next, for each topic we estimate a version of (4) in the CPS MORG data at the 10-year horizon, replacing our innovation exposure index $\eta_{i,t}$ with the topic k predictor. Finally, we form an in-sample data-mined displacement factor by taking linear combinations (either the average or the first principal component) of the candidate topics that are individually statistically significant negative predictors of employment declines at the 5% level.

Panels A and B of Table 2 compare the performance of our technology exposure measure (3) to the statistical labor-displacement factor. There are two points worth noting. First, these machine-learning predictors of employment declines also predict declines in occupation-level wages even though they are not explicitly constructed to do so. Second, and more importantly, these predictors perform about as well in predicting employment and wages as our baseline technology exposure measure $\eta_{i,t}$. This is not particularly surprising given that the correlation between our baseline measure $\eta_{i,t}$ and these machine-learning predictors is over 70 percent.

2.4 Comparison to existing measures of exposure to technical change

A key advantage of our measure relative to existing work aimed to measure workers' exposure to technological change (Autor and Dorn, 2013; Webb, 2019) is that it incorporates time-series variation by exploiting the arrival of breakthrough innovations. That said, it is instructive to explore the extent to which it contains additional information regarding cross-sectional differences in technology exposures. Here, we compare the performance of our technology measure in predicting wage and employment declines relative to the routine-task intensity measure from Acemoglu and Autor (2011),[14] 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). Since the Webb (2019) measures are calculated in percentile terms, we convert our index of technology exposure and routine-task intensity to cross-sectional percentile ranks to facilitate comparison of coefficients across measures. Our specifications take the form:

$$\frac{100}{h}\Big(\log Y_{i,t+h} - \log Y_{i,t}\Big) = \alpha + \alpha_j + \beta \eta_{i,1980}^{\text{Pctile}} + \delta X_i^{\text{Pctile}} + \epsilon_{i,j} \tag{5}$$

In estimating (5), 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. Here i indexes occupations and j indexes

[14] The construction of routine-task intensity in Acemoglu and Autor (2011) uses O*NET occupational information. Similar results obtain if we alternatively use the routine-task intensity index from Autor et al. (2003), which is constructed from the DOT.

industries. We include industry fixed effects α_j to account for industry specific shocks that may be correlated with occupational outcomes. We use the 1980 start-of-period percentile rank of our exposure measure. Depending on the specification, the variables X_i include the routine-task intensity percentile rank or Webb (2019) measures of exposure to robotics or software patents. We weight observations by the employment share in 1980 and cluster standard errors by occupation.

Tables A.4 reports our findings for employment (Panel A) and wages (Panel B). We estimate versions of (5) with our exposure measure by itself, with and without industry fixed effects, and then include the Webb (2019) and Acemoglu and Autor (2011) measures individually; finally, we estimate A.4 with all four measures simultaneously. We see that the point estimates of β are negative and highly significant, and they are largely unaffected by including the Acemoglu and Autor (2011) or Webb (2019) measures. We conclude that cross-sectional differences in our measure of occupational exposure to technology contain independent information relative to these alternate cross-sectional metrics.

3 Descriptive Patterns in Technology Exposure

Here, we document the workers' technology exposure has varied over time and across occupations.

3.1 Which Occupations Are More Exposed?

We find significant differences in the average degree of technology exposure across occupations. Over the entire sample starting in 1850, we find that the most exposed 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. By contrast, service-type occupations that specialize in person-to-person interaction score especially low on average exposure. Table A.3 lists the top and bottom 25 occupations by average exposure over the entire 1850-2002 sample period for $\eta_{i,t}$.

We next examine how the average level of occupation technology exposure is related to occupation-level wage earnings. Figure 3 plots the occupation technology exposures against average wage percentile ranks for the post-1980 period. We obtain information on average wages by occupation from the Current Population Survey Merged Outgoing Rotation Groups, which provides data on wages for the post-1980 period. Given the short time dimension of the data (MORG starts in 1980), we focus on cross-sectional comparisons.

Examining Figure 3, we see that the most exposed occupations tend to be found in the middle of the income distribution. This pattern is consistent with the prevailing view regarding job polarization in the United States, namely the decline in employment and wages for workers in middle-wage occupations (Autor and Dorn, 2013; Bárány and Siegel, 2018). In particular, Autor and Dorn (2013) argue that this pattern of job polarization—the disappearance of middle-skill (wage) occupations—is driven by their increased exposure to technological innovation relative to low-skill occupations. Our direct measure of technology exposure confirms this view.

3.2 Long-Run Shifts in Highly Exposed Occupations

We next examine how the types of occupation that are most exposed to technological innovation 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 A.1 plots these shares over the entire sample (1850 to 2002).

Examining Figure A.1, 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.

Importance of Manual vs Cognitive Tasks

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.[15] 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

[15] 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.

technological exposure of task category w as follows:

$$\lambda_{w,t} = \sum_i \eta_{i,t} \times T_w(i) \times \omega_i \tag{6}$$

Panel A of Figure 2 plots the composition of our measure of technology exposure $\lambda_{w,t}$ across each of these task categories. 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 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.

Educational Requirements

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 a given year. For this analysis we crosswalk SOC 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 $\lambda_{w,t}$:

$$\zeta_{s,t} = \sum_i \eta_{i,t} \times S_{s,t}(i) \times \omega_{i,t} \tag{7}$$

Panel B of Figure 2 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. Here, it's 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.

4 Technology Exposure and Worker Outcomes

Armed with a measure of workers' technology exposure, we next move to the main goal of the paper: understanding how exposure to major technologies shapes worker outcomes. The availability of administrative data on the earnings of individual workers, combined with demographic information such as age, education or past earnings, allows us to not only understand the experience of individual workers but also understand how it varies in the cross-section.

4.1 Data

We use a random sample of individual workers tracked by the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) 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 within the past 3 years so that the occupation information is relatively recent.[16]

Our main outcome variable is the worker's earnings growth over the next h years. To smooth out the impact of transitory earnings spikes (for example, a bonus), we construct a measure of

[16] Our sample includes individuals coming from the 1981-1991, 1994, and 1996-2016 ASEC waves for whom valid individual identifiers (a Census Protected Individual Key) can be assigned. In selecting which years to include in our sample, we exactly follow the labor force attachment restrictions imposed by Braxton, Herkenhoff, Rothbaum, and Schmidt (2021); we refer the reader to their paper for more details.

(smoothed) wage earnings growth following Autor, Dorn, Hanson, and Song (2014); Guvenen, Ozkan, and Song (2014); Kogan et al. (2020)

$$g_{i,t:t+h} \equiv w^i_{t+1,t+h} - w^i_{t-2,t}. \tag{8}$$

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

$$w^i_{t,t+h} \equiv \log\left(\frac{\sum_{j=0}^{h} \text{W-2 earnings}_{i,t+j}}{\sum_{j=0}^{k} D(\text{age}_{i,t+j})}\right). \tag{9}$$

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

The administrative data has detailed information on the industry of a particular worker, and also provides the industry of origination of individual patents. This allows us to exploit additional sources of variation, and now compare industries not only across occupations in the same industry, but also workers in the same occupation across industries. To this end, we build our technology exposure measure by also restricting attention to patents issued to firms in the same 4-digit NAICS industry as the worker. Letting j index patents as before; $\mathcal{B}_{k,t}$ 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 worker i in year t, we measure the technology exposure of worker i to technology at time t as

$$\xi_{i,t} = \log\left(1 + \sum_{j \in \mathcal{B}_{k(i,t),t}} \rho_{o(i),j}\right). \tag{10}$$

In brief, our technology exposure metric (10) bears a strong similarity to (3), with the following modifications: $\xi_{i,t}$ defined in (10) now varies by occupation, industry, and year instead of just occupation and year; given that the resulting measure is quite sparse—just under half of the industry–occupation pairs have zero breakthrough patents in a given year—we apply a log transform to smooth out the resulting skewed distribution; and given that we are now considering a much shorter time series in these worker-level regressions we no longer need to express our measure in units of per capita patenting. Last, given that the administrative data sample has a shorter time dimension, we use the breakthrough definition of KPST that relies on 5-year (as opposed to 10-year) forward similarity, which allows us to extend the sample period for $\xi_{i,t}$ by five additional years up through 2007.

Table 3 summarizes the sample. After restricting attention to worker-year observations following the five year period that the worker appears in the CPS, we are left with approximately 2.8

million person-year observations spanning the period from 1988 to 2016. In terms of demographics, approximately 54% of the sample is male and 34% of the observations correspond to workers with a four-year college degree. The median worker in the sample is 41 years old and earns approximately $50k per year in terms of 2015 dollars. The distribution of earnings is rather skewed: the average is equal to $66k while the 5th and 95th percentiles are equal to $16k and $152k, respectively. The last three rows report the distribution of our the lifecycle-adjusted cumulative earnings growth (8) measure. At a horizon of $h = 5$ years, the median is equal to approximately zero while the mean is -0.095; given that (8) 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 10th percentile is equal to -0.61 log points while the 90th percentile is equal to 0.574.

4.2 Technology exposure and worker earnings growth

We estimate the following specification,

$$g_{i,t:t+h} = \alpha + \beta\,\xi_{i,t} + \delta Z_{i,t} + \varepsilon_{i,t}. \tag{11}$$

The dependent variable defined in (8) is the growth in worker i's average earnings over the next $h = 3, 5$ and 10 years, relative to the prior three years. The main dependent variable of interest $\xi_{i,t}$ captures the worker's exposure to breakthrough technologies in her (NAICS 4-digit) industry. The vector Z includes a rich set of controls that aim to soak up ex-ante worker differences. Depending on the specification, we include various combinations of year, occupation and industry fixed effects. Our most conservative specification, and the one we focus most in the paper, 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.[17] Standard errors are clustered at the industry level.

Table 4 reports the estimated slope coefficient β. Overall, we find that workers' technology exposure is negatively related to their subsequent earnings growth. Panel A reports the estimated slope coefficients β for horizons of $h = 3, 5$ and 10 years; different columns correspond to different fixed effect combinations. The magnitudes are both economically and statistically significant. Our preferred specification focuses on the 5-year horizon and the most conservative specification that

[17] 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.

compares a worker to, either other workers in the same industry but different occupation, or workers in the same occupation but different industries. In this case, we see that a one standard deviation increase in innovation is associated with a 0.02 log point decline in average worker earnings over the next five years. These magnitudes increase with the horizon h, ranging from 0.017 to a 0.02 log point decline in average earnings at horizons of three and ten years, respectively. Comparing across columns, we see that the coefficient estimates mostly increase in magnitudes as we saturate the specification with additional controls, suggesting that the relation we identify is driven by primarily within-industry and within-occupation variation. Accordingly, for the rest of the paper we focus our attention on the specification corresponding to column (4) of Table 4.

The negative average correlation between technology exposure and worker earnings can potentially mask considerable heterogeneity worker outcomes across job types. To that end, we next allow the estimated slope coefficient to vary across samples. Panel B compares the estimated coefficient β across workers employed in manufacturing and services, broadly defined. We find that the estimated coefficient is essentially the same in both sectors—that is, workers in both manufacturing and services experience roughly a similar decline in earnings in response to a shock to ξ. This pattern, combined with the increase in the average technology exposure of workers in the service sector, relative to manufacturing we saw in Section 3.1, strongly suggests that the displacement of workers in response to labor-saving technologies is not purely a blue-collar worker phenomenon, but it is (increasingly) present in workers employed in service industries as well.

Digging deeper, we next allow the slope coefficient β to vary across different types of jobs—specifically, across occupations that score highly in the Acemoglu and Autor (2011) categories. Table 5 reports how the estimated coefficient varies across occupations that are above or below the median in the categories: manual physical; non-routine manual and personal; routine or non-routine cognitive tasks. As we examine the table we see that the negative relation between our technology exposure measure ξ and subsequent earnings growth is present across most occupations—with the notable exception of occupations that score highly in terms of non-routine manual and interpersonal skills. That said, the magnitudes are not always the same: the negative relation we document is particularly salient for workers in occupations that emphasize manual and routine cognitive tasks—exactly the same occupations that have been exposed to most of the technological breakthroughs as we saw in the top panel of Figure 2.

4.3 Worker 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 and prior income as common proxies for worker skill; further, allowing the impact of technology to vary by worker age helps tease out the effect of vintage-specific human capital.

Education

Panel A of Table 6 compares the estimated coefficient across workers with and without a college degree. Perhaps surprisingly, and somewhat inconsistent with the standard view of technology-skill complementarity, we find no meaningful differences in the response of worker earnings to ξ among workers with and without a college degree. Put differently, regardless of whether they have a four-year college degree or not, two workers in the same industry and same occupation will on average experience the same decline in wage earnings in response to the same increase in technology exposure.

Age

The bottom panel of Table 6 shows that older workers (those in the 45 to 55 range) experience significantly greater declines in earnings growth (0.02 to 0.026 log points across horizons) relative to workers aged 35–45 (0.012 to 0.019 log point decline) or 25–35 (0.008 to 0.014 log point decline). This strong gradient within age is consistent with the view that 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.

Income

We next examine how the response of worker earnings growth to our technology exposure measure $\xi_{i,t}$ varies by the worker's prior income relative to her peers. All else equal, one would expect that workers that are paid more relative to their peers in the same industry and occupation have accumulated a greater amount of skill at performing the tasks required by their occupation. A common view of technology is that it is complementary to worker skill, and as such we might expect that the most highly workers respond less—or even experience wage increases—as the technology that is relevant to their tasks improves.

Our findings suggest a more nuanced view. Column (1) of Table 7, shows how the slope coefficient β to vary with the worker's current income relative to her peers in the same industry–occupation cell. We see that workers at the top of the relative earning distribution experience significant greater wage earnings declines (0.042 log points) that is more than twice the magnitude for the average worker (0.02 log points) in response to a unit standard deviation increase in technology exposure $\xi_{i,t}$. Columns (2) to (4) show that this pattern is quite distinct from the age pattern we documented above. Specifically, when we rank workers in terms of their prior income relative to other workers in the same age group within their industry–occupation cell, the gradient on income remains largely similar: the estimated slope coefficient β is approximately twice as large in terms of magnitude for the highest-paid workers for each age group.

One interpretation of these patterns is that part of the worker's accumulated skills (human capital) is vintage-specific. As technology improves, part of these skills become obsolete. Workers who have accumulated the most skill in the existing technology also have the most to lose when new technologies arrive. However, an alternative interpretation is that the adoption of labor-saving technologies is more likely in certain jobs or firms where workers are paid 'excessively' high wages relative to their peers—jobs in which employees have greater bargaining power and therefore can appropriate a larger share of the value added. The next set of facts suggests that this alternative mechanism is unlikely to fully account for this income gradient.

Appendix Table A.6 explores two alternative approaches for constructing income rankings that aim to at least partially address this concern. First, we adjust the worker's earning rank for average wage differences across firms, since we don't always have a sufficient number of workers at each firm to compute a within-firm earnings rank for each worker. We do so by ranking workers based on their earnings relative to the firm's average wage—more specifically, total wage bill divided by employment from the LBD. Second, we rank workers based on a wage residual from a regression aimed to remove heterogeneity in worker earnings unrelated to worker skill within an occupation–industry. Specifically, in addition to removing yearly occupation- and industry-fixed effects from log wages, we remove fixed effects for unionization status (when available); gender interacted with 10-year age bins; and worker commuting zones, all interacted with calendar year. In particular, removing commuting zone (interacted with year) effects allows us to remove the component of pay that may be related to local labor scarcity or monopsony power, while removing age bin by gender fixed effects accounts for the fact that older workers tend to be more highly paid. Columns (3) and (4) show two alternative residualizations, where column (4) focuses on the subsample of about one-fifth of workers where who were in-universe for the unionization question on the CPS. As we compare column (1) to columns (2), (3), and (4), we see that our results remain largely unchanged: the highest paid workers given these alternate definitions still experience greater wage earnings declines in response to an increase in technology exposure.

Appendix Table A.7 explores the extent to which the income gradient varies with the level of unionization—a proxy for increased worker bargaining power. In columns (1) and (2), we see that the income gradient we document is present in industries with both high and low levels of unionization, though the magnitudes are slightly larger for workers in industries with low unionization rates. More importantly, columns (3) and (4) allow the effect to vary depending on the worker's own union membership status. Even though the sample is dramatically smaller (we have unionization information for approximately one-fifth of the workers in our sample), we again find no meaningful differences in the slope of the income gradient as a function of union membership.

Overall, this evidence suggests that the larger earnings responses to technological exposure among more highly paid workers are not being driven by targeted technology adoption to replace workers that are 'overpaid' relative to their peers (e.g. due to greater bargaining power). Instead, our findings are more consistent with skill displacement, where highly productive workers see steeper relative earnings declines following the introduction of new technologies that substitute for the current vintage of tasks in which they specialize.

4.4 Technology exposure and worker earnings risk

As we discussed above, one potential reconciliation of the patterns in Table 7 with the standard view of technology-skill complementarity is that part of human capital is specific to a particular vintage. Put differently, skill is not an immutable characteristic of the worker; it is the result of experience and learning by operating a particular technology. As 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.

We next examine the degree to which the differences in earnings responses to ξ are driven by increased likelihood of large earnings declines—that is, movements in the left tail as opposed to the conditional mean. To this end we re-estimate equation (11), where now the dependent variable is a dummy that takes the value one if the worker's earnings growth over the next five years is below the 10-th percentile of the earnings growth distribution across all workers in the same year.

Examining Table 8, we see that top earners face significantly greater labor income risk than the average worker in response to an increase in their technology exposure. Focusing on our preferred specification, a one-standard deviation increase in $\eta_{i,t}$ is associated with a 1.62 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. In terms of magnitudes, this increase in tail risk accounts for approximately one-third of the negative mean effect.

We next explore the underlying drivers of this increase in tail risk. One possibility is that more highly-paid workers face increased likelihood of unemployment in response to improvements in technology related to their tasks. Appendix Table A.8 shows that this is unlikely to be the main driver for this increase in risk. Even though the likelihood of unemployment—defined either as one year with no W2 earnings in column (2) or three consecutive years of zero W2 income in column (3)—does increase for workers across all income groups, there is no meaningful difference in the estimated slope coefficients across income groups. Similarly, Appendix Table A.9 shows that the differences in wage responses as a function of income are essentially the same if we condition the sample to workers that have not experienced unemployment.

4.5 Heterogeneity as a function of related experience

A direct implication of our vintage human capital interpretation of these results is that the income pattern we document should be stronger in occupations in which human capital is both important and specific. To this end, we use data on occupational related work experience requirements from the 2010 O*NET to proxy for the amount of specific skills needed for a particular job. For each occupation O*NET reports occupational related work requirements in time intervals—such as 1 to 3 months, 1 to 2 years, 10 years or more, etc.—as well as corresponding weights for how frequently a particular interval is reported for that occupation. We compute occupational expected years of related experience by taking the middle of each interval in year terms (with the exception of the highest bin, which we top code at 10 years), then take a weighted average using the frequency weights. Finally, we average the resulting expected years of related experience across all O*NET SOC codes that are crosswalked to a given occ1990dd code. We then re-estimate equation (11) while also allowing the slope coefficient to vary across occupations that score above or below the worker-level median in terms of required related experience.

Table 9 presents our findings. Focusing on the first two columns, we see that the income gradient in the estimated slope coefficients is significantly higher for occupations that require longer periods of related experience. In occupations in which related experience is important, workers in the highest-paid workers experience a 0.044 log point decline in their cumulative earnings over the next five years in response to a unit standard deviation increase in ξ, compared to 0.016 log point decline for the lowest-paid workers. By contrast, difference between the highest- and lowest-paid group in the occupations in which related experience is less important is -0.037 and 0.029, respectively.

The next two columns show that the income gradient in tail risk is also stronger in occupations in which related experience is required. In these occupations, workers in the highest-paid group experience a 1.63 percentage point increase in the likelihood of a large earnings decline in response to a unit standard deviation increase in ξ, compared to a 0.50 percentage point increase for the lowest-paid group. In occupations in which related experience is less important, the corresponding estimates are 1.32 and 1.28 percentage points, respectively.

4.6 Is all technology labor-saving?

Our findings so far seem to suggest that there is a consistently negative effect of technology exposure on employment and wage growth. It is important to keep in mind, however, that this conclusion is conditional on the construction of our technology exposure measure, which is effectively a count of breakthrough patents (in the same industry) that are similar to tasks performed by each occupation. Inasmuch as task descriptions' textual overlap with technologies captures task substitution, as we have suggested, this measure isolates the subset of innovations that are most likely to displace workers.

In what follows, we perform the reverse exercise and identify the breakthrough patents issued to firm's in the worker's industry, that are however *least likely* to be related to the tasks performed by a given worker. We construct this exposure measure in a similar fashion as (10),

$$\zeta_{i,t} = \log\left(1 + \sum_{j \in \mathcal{B}_{k(i,t),t}} d_{o(i),j}\right). \tag{12}$$

The key difference is that now $d_{o(i),j}$ is a measure of distance rather than similarity between the description of patent j and the tasks performed by occupation o. We construct d in an equivalent fashion as ρ except that our starting point is $1 - Sim_{i,j}$ from equation (2) instead of $Sim_{i,j}$. We then remove patent issue year fixed effects from $1 - Sim_{i,j}$ as before. Finally we follow our previous procedure by setting all pairs beneath the 80th percentile of fixed effect-adjusted of textual dissimilarity equal to zero, and scale scores between the 80th and 100th percentiles such that the 80th percentile equals zero and the maximum equals one. Just as before, this measure gives more positive weight to patents with textual "relatedness" sufficiently far in the right tail of the distribution; however, now our notion of "relatedness" is textual dissimilarity rather than textual similarity. The objective here is to capture patents whose textual descriptions are unrelated to an occupation's task descriptions; based on our findings thus far which suggest that textually similar patents predict displacement of labor, we expect this measure to now capture patents that are especially unlikely to displace the skills of a given worker. We therefore test whether $\zeta_{i,t}$ has a positive effect on worker wages.

We next examine the impact of high- and low-task similarity of breakthrough technologies on worker earnings. To do so, we estimate the following variant of (11)

$$g_{i,t:t+h} = \alpha + \beta\,\xi_{i,t} + \gamma\,\zeta_{i,t} + \delta Z_{i,t} + \varepsilon_{i,t}. \tag{13}$$

Table 10 reports the estimated coefficients β and γ from Equation (13). For brevity we focus on our preferred specification, which is saturated with industry times year and occupation times year fixed effects, though results are similar with less saturated specifications.

The top panel of Table 10 reports results for all workers; the bottom panel allows the estimated coefficients β and γ to vary with the worker's prior income as before. There are three things to note. First, the estimated coefficient γ is positive and statistically significant: the average worker experiences approximately a 0.01 log point increase in her average earnings in response to a unit standard deviation increase in ζ. Second, with the exception of the workers in the bottom-25th percentile, there is again an income gradient in terms of responses of wage earnings to ζ: workers at the top of the earnings distribution experience approximately a 0.014 increase in their cumulative earnings over the next five years, which is approximately 50 percent larger than the response of the average worker in the 25th to 95th percentile (the difference is statistically significant at the 10 percent level). Third, including ζ in this specification has very little influence on the estimated coefficients β relative to the specification in (11): a one-standard deviation increase in ξ is associated with approximately a 0.015 log point decline in worker earnings and, as before, the earnings of top workers respond more than the average worker.

In sum, these findings reinforce our view that our measure ξ primarily captures technological breakthroughs that are labor saving. Technological improvements that are unrelated to a worker's task are in general associated with an increase in subsequent earnings growth. Importantly, with the exception of the very lowest-paid workers, the earnings of the workers at the top of the income distribution respond more to an increase in ζ than the average worker.

4.7 Discussion

Taken together, the findings in the previous sections indicate that improvements in technology that are closely related to the tasks performed by a given worker are associated with lower future earnings. This pattern is distinct from breakthrough technological improvements in the same industry that have low similarity with the tasks performed by the worker. We find that the decline in earnings is significantly higher for workers that are more highly paid relative to their peers in the same occupation and industry. Importantly, this income gradient is steepest in occupations for which related experience is more important—a proxy for specific human capital.

To the extent that workers in the highest-paid group receive higher wages due to their higher ability or skill, this pattern can be surprising under a somewhat naive view of technology-skill complementarity: if skill is an immutable characteristic of the worker, we would expect to see that more highly-skilled workers experience either lower wage declines or wage increases as they are more likely to perform tasks that are complementary to the underlying technology, or equivalently, the tasks they perform are less likely to be automated because they are more complex and command higher pay. In the next section we develop a simple model in which technology increases the skill premium yet highly paid workers experience a greater risk of income declines.

We find little evidence that this income gradient varies across worker groups with varying degrees of bargaining power (unions). Also, we obtain similar patterns when we rank workers on income net of several observables including unionization status or commuting zone fixed effects. Both patterns suggest that selective adoption of labor-saving technologies to the highest paid workers—though certainly plausible—are not the key driver behind the income gradient we document.

5 Model

The increased earnings response of more highly-paid and older workers suggests a role of vintage-specific human capital—technology making certain worker skills obsolete. To this end, we develop a model in which workers perform two tasks, one of which is a complement and another which is a substitute to technology. The complementary task is more highly rewarded and hence skilled workers perform tasks that are less likely to be automated and instead are complements to technology. However, 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.

5.1 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

$$Y_t = \left[\mu\Big(H_t\Big)^{\sigma} + (1-\mu)\left(\lambda\Big(\xi_t\Big)^{\rho} + (1-\lambda)\Big(L_t\Big)^{\rho}\right)^{\sigma/\rho}\right]^{1/\sigma} \tag{14}$$

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. Since the total mass of workers is normalized to one, equation (14) also refers to labor productivity (output per worker).

The factor ξ is the stock of intangible capital/knowledge embodying the technology used for producing output Y_t, similar in spirit to Acemoglu and Restrepo (2018). We allow technology to be more complementary to skilled labor relative than 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

$$\sigma < \rho < 1. \tag{15}$$

Put differently, our notion of skill pertains to worker's ability to use technology as a complement in production. Technology ξ evolves exogenously according to

$$d\xi_t = -g\,\xi_t\,dt + \kappa\,d\,N_t. \tag{16}$$

Technology improves according to the Poisson process N_t with arrival rate $\omega\,dt$. Recall that we have set up output in per capita terms. As such, the negative drift term in equation (16) reflects the fact that ξ also a per-capita quantity and population grows at rate g. Given (16), the level of ξ is stationary with a long-run mean equal to $\kappa\,\omega/g$.

When we map our empirical analysis to the model, we will interpret our empirical I technology exposure metric $\xi_{i,t}$ as a shock to ξ, which however affects only a subset of workers involved in the production of Y—the model equivalent to 'occupations'. Specifically, 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

$$H_t = \int_0^1 \theta\, p_t(\theta)\, d\theta, \tag{17}$$

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,

$$L_t = 1 - H_t. \tag{18}$$

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 $\theta \in (\underline{\theta}, 1)$. The skill level of workers with $\theta \in [\underline{\theta}, 1]$ evolves over time due to learning by doing and technological displacement according to the process

$$\frac{d\theta_{i,t}}{\theta_{i,t}} = m_{i,t}\,dM_{i,t} - h\,dN_{i,t}, \tag{19}$$

which we further assume to be reflected at the boundaries of the interval $[\underline{\theta}, 1]$.[18] In the above

[18] Reflecting the process at the boundaries of the interval guarantees that the measure of workers who can acquire new skills, s_h, is well defined and constant over time. Specifically, if following a realization of the jump $dM_{i,t}$ the process $\theta_{i,t}$ were to exceed 1 and thus escape the interval $[\underline{\theta}, 1]$, we set the value of $\theta_{i,t}$ immediately following the jump to 1. Similarly, if a realization of $d\,N_{i,t}$ were to lead $\theta_{i,t}$ to fall below $\underline{\theta}$, we set the process immediately after the jump to $\underline{\theta}$.

specification, $M_{i,t}$ is a Poisson process, independent across the workers, with arrival rate ϕ which governs learning/forgetting which is independent of technology. Specifically, conditional on $dM_{i,t} = 1$, a second random variable $m_{i,t}$ governs whether a worker learns or forgets; skills grow by $m\%$ with probability $1 - z$ and shrink by $m\%$ with probability z. This formulation captures the idea that acquiring new expertise is an uncertain process, and the parameter z allows for a baseline level of risk that skills depreciate.

Importantly, the last term in equation (19) captures the displacive effect of the arrival of new technologies which is stochastic. The process $N_{i,t}$ is a doubly stochastic jump process driven by the Poisson process N_t, which describes technological improvements in (16). Specifically,

$$dN_{i,t} = d_{i,t}\, dN_t. \tag{20}$$

Each technological improvement affects workers randomly, lowering their skill level θ by a random magnitude driven by $d_{i,t}$, which has a support on the unit interval, is independent of $\theta_{i,t}$, and independently distributed across workers. For now, we assume that $d_{i,t}$ follows a binomial distribution $d_{i,t} \in \{0, 1\}$ with $Prob(d_{i,t} = 1) = \alpha$. More generally, we could allow the distribution of $d_{i,t}$ 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. Finally, workers of each type die at Poisson rate δ 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 s_h, respectively. 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.

Given our assumptions (17)–(20), 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

$$w_{i,t} = W_{L,t} + \theta_{i,t} \left(W_{H,t} - W_{L,t} \right). \tag{21}$$

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

$$W_{H,t} = \frac{\partial Y_t}{\partial H_t}, \quad \text{and} \quad W_{L,t} = \frac{\partial Y_t}{\partial L_t}. \tag{22}$$

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.

5.2 Model Calibration

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

Methodology

The model has a total of 15 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. Table 12 summarizes the 13 statistics that we target to calibrate the remaining 11 parameters $\Theta = \{\mu, \lambda, \rho, \sigma, \phi, \alpha, \kappa, \omega, h, g, z\}$.

We target the average skill premium in the data, defined as the mean ratio of earnings of workers in the 75th vs the 25th percentile within an industry. 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, the mean skill premium affects the parameters driving the ergodic distribution of θ, namely ω, ϕ, h, z, and α.

We calibrate the model impulse response to ξ of labor productivity (14) and the labor share to the data. In the model, an increase in ξ has an ambiguous impact on both output and the labor share. It depends 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—recall that the aggregate supply of high- and low-skill inputs H and L varies in the short run, due to skill displacement (19).

To obtain an empirical analogue to the response of labor productivity and the labor share to technology, we rely on (and extend) the analysis in Kelly et al. (2021). In brief, we obtain data on industry-level measures of output per worker and the labor share from the NBER manufacturing database–which cover the 1958 to 2018 period. We assign breakthrough patents to industries based on their CPC technology class using the probabilistic mapping constructed by Goldschlag, Lybbert, and Zolas (2020). We then estimate the response of productivity and the labor share over the next five years. Appendix A.3 contains more details.

We find that a one-standard deviation improvement in technology is associated with a 0.0281 log point increase in industry labor productivity and a 0.0125 log point decline in the labor share over the next five years. 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. The fact that output/productivity and the labor share respond with opposite signs helps narrow down the set of admissible parameters quite significantly.

To help identify the parameters involved in the dynamics of worker skill acquisition and displacement in (19), we also target the heterogeneity in earnings responses to changes in technology (see Section 4). Specifically, we target the mean earnings growth responses and changes in worker tail risk—column (1) in Table 7 and column (4) in Table 8. In the model, whether higher-paid workers are more exposed to technology than lower-paid workers is largely ambiguous. To see this, we decompose wage earnings growth in the model over the next h periods,

$$\frac{w_{i,t+h} - w_{i,t}}{w_{i,t}} = \underbrace{\frac{w_{l,t}}{w_{l,t} + \theta_{i,t} s_{p,t}}}_{\substack{\text{low skill} \\ \text{income share}}} \underbrace{\frac{\Delta_h w_{l,t+h}}{w_{l,t}}}_{\substack{\text{low skill} \\ \text{wage chg}}} + \underbrace{\frac{\theta_{i,t} s_t}{w_{l,t} + \theta_{i,t} s_t}}_{\substack{\text{high skill} \\ \text{income share}}} \left[\frac{s_{p,t+h}}{s_{p,t}} \cdot \underbrace{\frac{\Delta_h \theta_{i,t+h}}{\theta_{i,t}}}_{\substack{\text{skill} \\ \text{displacement}}} + \underbrace{\frac{\Delta_h s_{t+h}}{s_t}}_{\substack{\text{high skill} \\ \text{wage chg}}} \right] \tag{23}$$

where $s_t = W_{h,t} - W_{l,t}$ is the skill premium in differences.

As we see from the last term in brackets in (23), 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 θ due to skill displacement—see equation (19). For the high-income (high-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 shifts in skill prices following displacement.

Mapping the empirical regressions in Tables 7 and 8 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. 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.

We choose $\Theta = \{\mu, \lambda, \rho, \sigma, \phi, \alpha, \kappa, \omega, h, g, z\}$ by minimizing the distance between the output of the model $\hat{X}(\Theta)$ and the data X,

$$\hat{\Theta} = \arg\min_{\Theta} \left(X - \hat{X}(\Theta)\right)' W \left(X - \hat{X}(\Theta)\right). \tag{24}$$

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 11 summarizes our parameter choices. Similar to Krusell et al. (2000); Eisfeldt, Falato, and Xiaolan (2021), we find that technology is a good substitute for the low-skill labor input ($\rho = 0.68$) whereas the high-skill labor input is complementary to technology ($\sigma = -0.04$). Technology shocks are relatively frequent ($\omega = 2.06$) and sizeable ($\kappa = 0.16$). 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 skill θ. That said, these losses are pervasive (the probability of skill loss conditional on a shock is $\alpha = 23\%$) though transient: workers are able to acquire skills (increase θ) at an average rate of $m\,\phi\,(1 - 2z) = 5.8\%$ per year.

Model Fit and Discussion of the Mechanism

Examining Table 12, we see that the model does a good job matching the 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 patterns in the marginal effects of shocks on exposed workers in terms of income growth rates and left tail risk.

Figure 5 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 2.7% rise in output/productivity on impact. By contrast, Panel C shows that the labor share declines by approximately 1.2%. This decline in the labor share in the model 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% as workers' skills are displaced. This fall is temporary, as H gradually increases over time as workers acquire more skill. 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 a 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 6 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. The closest empirical analogue to these responds is the second column in Table 10. Comparing these numbers to their model analogues, we see a qualitatively similar pattern (with the exception of the workers in the bottom-25 percent) but the income gradient for the low-exposed workers is stronger in the model than the data.

The orange bars in Panel A of Figure 6 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 6 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 non-monotonic response, with the workers in the lowest income bin experiencing a slightly larger drop in earnings than those in the second bin.

In brief, Figure 6 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 a steep income gradient in earnings losses following increases in technology. 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.

The race between education and technology

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. Figure 7 displays the results of two experiments. The blue line corresponds to transition paths associated with a permanent increase in the rate of technological innovation (ω) 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. Panel B shows that a permanent increase in the level of ω leads to a permanent decline in the quantity of the supply of the high-skill input due to a continual rate of displacement. In response to higher ω we see an increase in output and productivity (Panel C), which is somewhat transitory due to the fact that H declines and H and ξ are complements; a decline in the labor share (Panel D), and an increase in the skill premium (Panel E). Importantly, even though the skill premium is higher in the new steady state, income inequality, as measured by the top 5 percent share, is actually lower. Panel F shows that even though income inequality initially rises, it is smaller in the new steady state: the high rate of displacement implies that few workers are able to acquire and retain a sufficiently high level of skill.

In the second experiment we also vary the rate of skill acquisition ϕ to keep the steady-state level of H constant between the two steady states. We can interpret this experiment as an increase in education aimed at maintaining the economy's current level of skill. In this case, output rises faster (since now H remains constant) while the labor share still falls. More importantly, however, we see that even though the increase in the skill premium now is much smaller than the previous experiment, income inequality in the new steady state is actually higher.

This exercise highlights the distinction between the skill premium and income inequality. In the model, these are distinct objects: income inequality depends not only on skill prices, but also the quantity of skill workers have. Changes in technology affect both quantities and prices, hence movements in the skill premium are insufficient to fully characterize earnings inequality.

6 Conclusion

We develop a methodology for identifying the arrival of labor-saving technologies that relies only on the textual description of the patent document and the tasks performed by workers in an occupation. Our measure primarily identifies labor-saving innovations and is broadly available from the 19th century to the present. 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.

Using administrative earnings records from the US Social Security Administration, we find that, in response to an increase in technology, the average worker experiences approximately a 0.02 log point decline in her wage earnings over the next five years. Importantly, we find significant heterogeneity in these responses across age and income levels. In particular, older workers are significantly more affected than younger workers. In addition, workers at the top end of the earnings distribution—relative to their peers in the same occupation and in the same industry—experience a significantly greater decline in earnings (more than twice) relative to the average worker. Our interpretation of these pattern is that it is consistent with the importance of vintage-specific human capital. That is, improvements in technologies are often associated with obsolescence of certain worker skills—or simply with a period of learning during which productivity is depressed. To the extent that workers that are more highly paid relative to their peers have acquired more skills, we would expect such workers to experience lower wage growth as their skills become obsolete. Consistent with this view, we find that the differential response of high-paid workers relative to the lower-paid group is significantly stronger in occupations that require a greater amount of related experience.

We rationalize our findings in the context of a model that features skill-biased technical change (Krusell et al., 2000) and vintage-specific human capital (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). The calibrated model quantitatively replicates the facts in the data. 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.

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

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

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

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

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

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

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

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.

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

Bessen, J., M. Goos, A. Salomons, and W. van den Berge (2022). What happens to workers at firms that automate? Working paper.

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

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

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

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.

Caunedo, J., D. Jaume, and E. Keller (2021). Occupational Exposure to Capital-Embodied Technical Change. CEPR Discussion Papers 15759, C.E.P.R. Discussion Papers.

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.

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

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

Deming, D. J. (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.

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

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature 57*(3), 535–74.

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.

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

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

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*.

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 (2021). Measuring technological innovation over the long run. *American Economic Review: Insights 3*(3), 303–20.

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

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.

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

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

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.

Figures and Tables

Figure 1: Occupations Most Related to Three Breakthrough Innovations

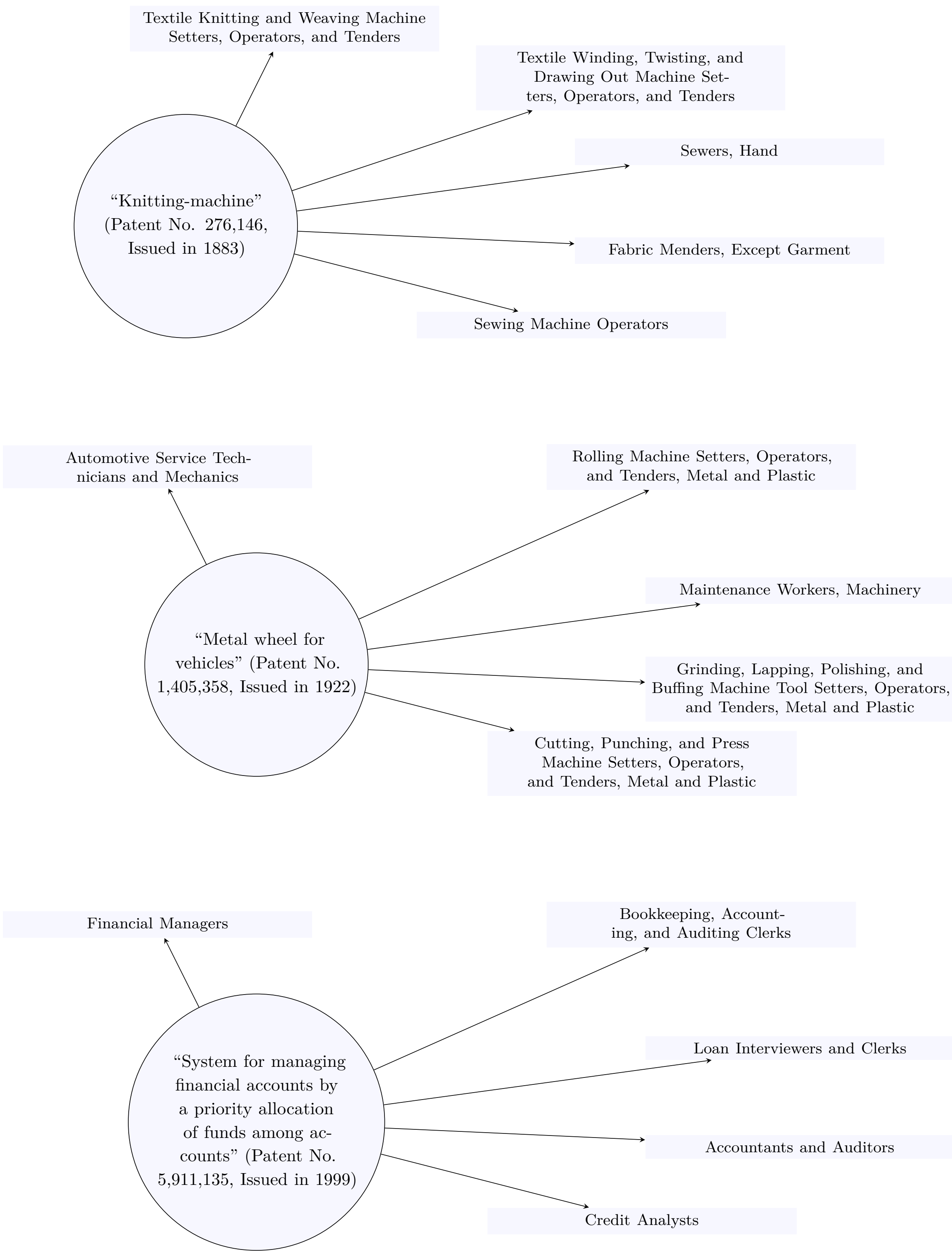
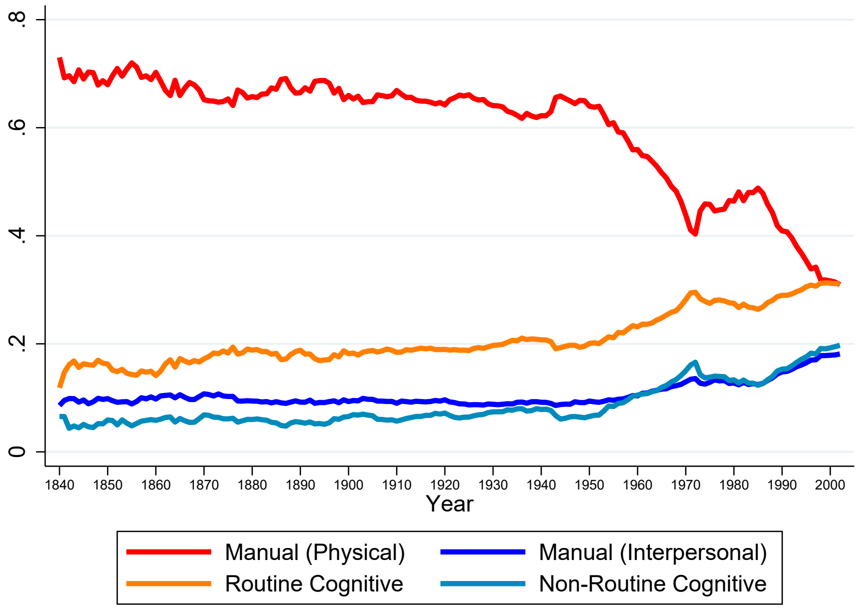
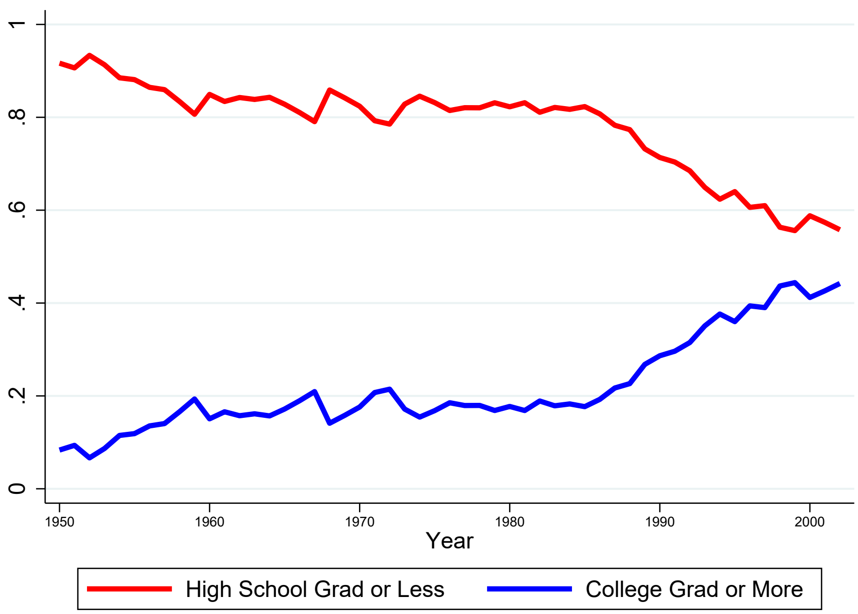


Figure 2: Technology exposure, composition across occupation categories

A. By Task Type

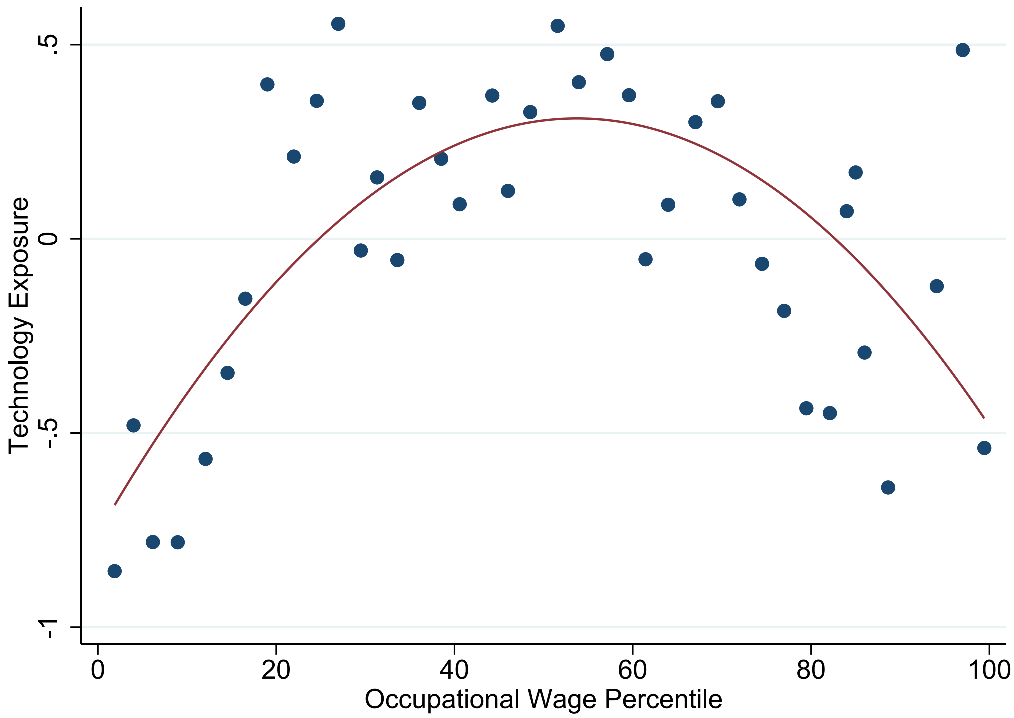


B. By Education Requirement



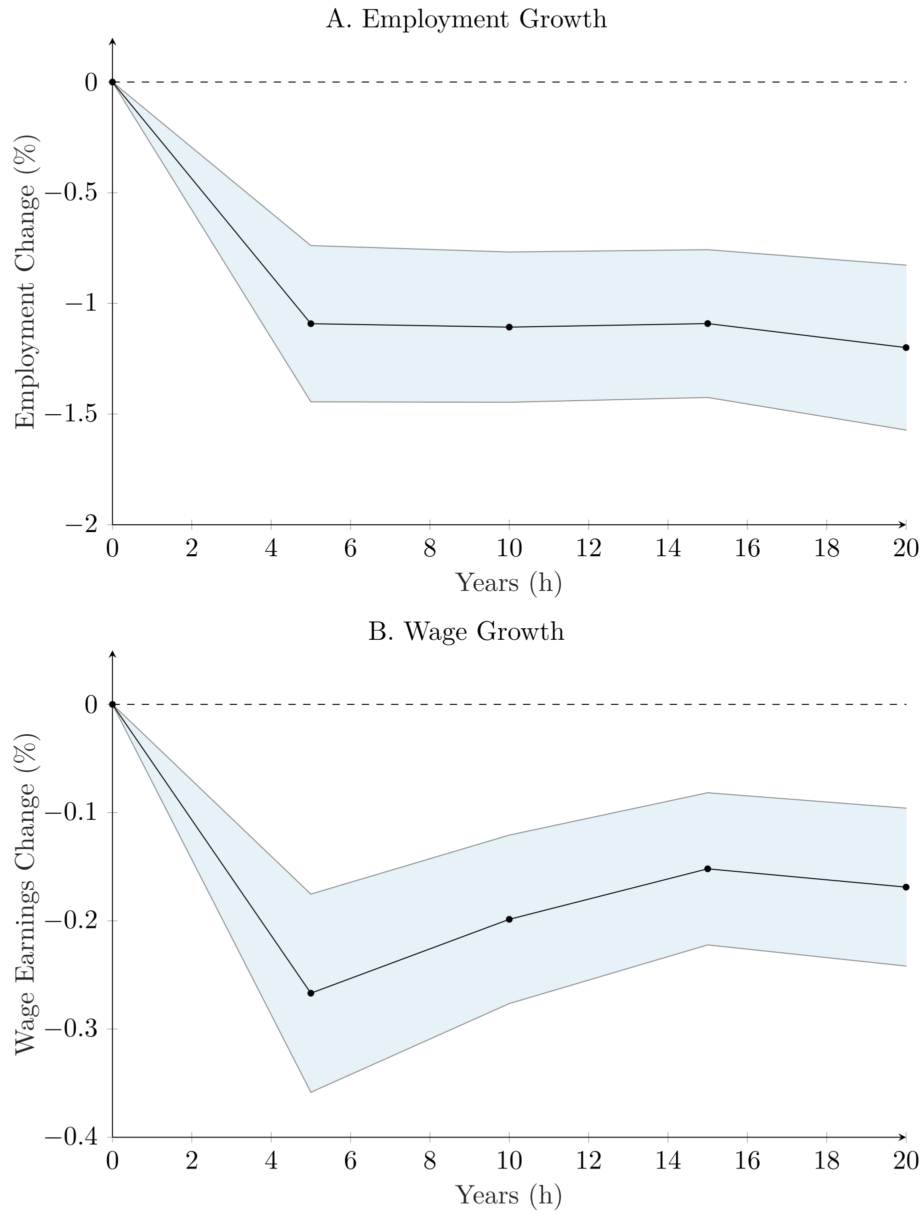
Note: Panel A of the figure plots the composition of our index of technological exposure by task category, $\lambda_{w,t} = \sum_i \eta_{i,t} \times T_{w,t}(i) \times \omega_i$. Here w represents one of the four given task categories. $T_{w,t}$ is an indicator that takes a value of 1 if occupation i 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. We plot the relative shares $\lambda_{w,t}/\sum_{w'} \lambda_{w',t}$. Panel B performs the analogous exercise by education requirements, $\zeta_{s,t} = \sum_i \eta_{i,t} \times S_{s,t}(i) \times \omega_{i,t}$, where now 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.

Figure 3: Technological Exposure, by occupation income level



Note: This figure plots a bin scatter for technology exposure $\eta_{i,t}$ by occupational wage percentile. The period covers the years from 1980 to 2002, and occupations are sorted by wage percentile rank. The wage data come from the Current Population Survey Merged Outgoing Rotation Groups.

Figure 4: Employment, wage earnings and technology exposure (recent period: 1980–present)

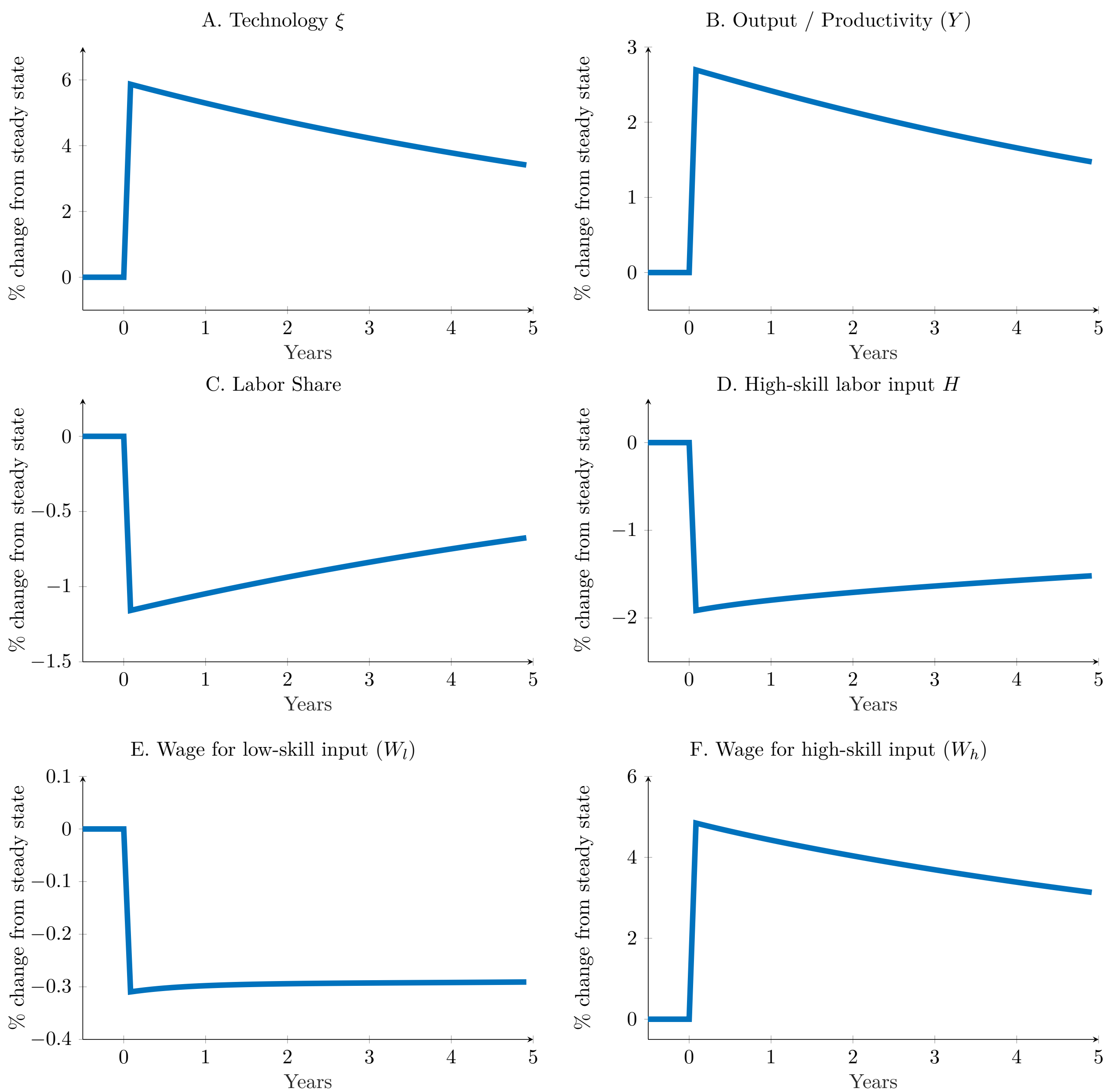


Note: The Figures above plot coefficients from panel regressions of annualized wage and income growth rates over different time horizons on occupation innovation exposures:

$$\frac{100}{h}\Big(\log Y_{i,t+h} - \log Y_{i,t}\Big) = \alpha + \beta\eta_{i,t} + \delta X_{i,t} + \varepsilon_{i,t}$$

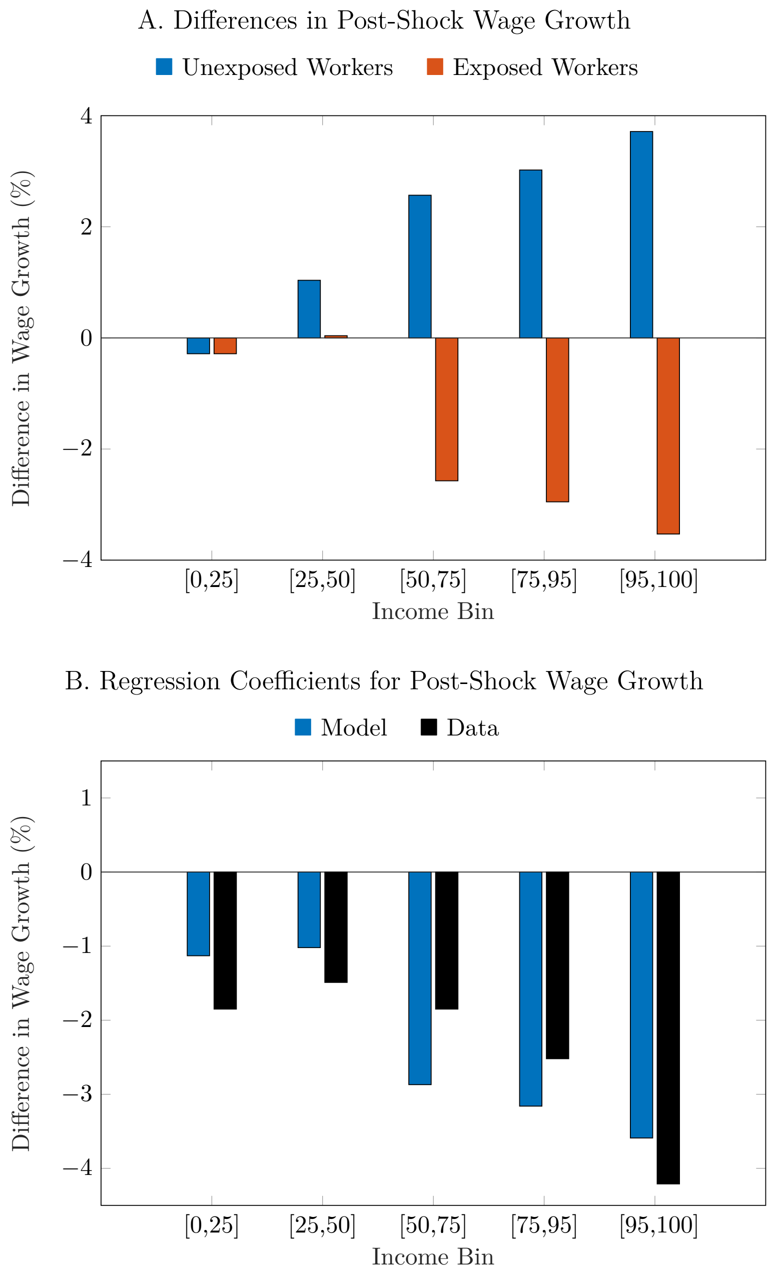
Here Y_i denotes occupation employment share or wage. Controls $X_{i,t}$ include three one-year lags of dependent variable, and time fixed effects. Dependent variable is expressed in annualized percentage terms and $\eta_{i,t}$ is standardized. Figures plot 90% confidence interval for each time horizon.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

Figure 5: Model: Impulse Responses



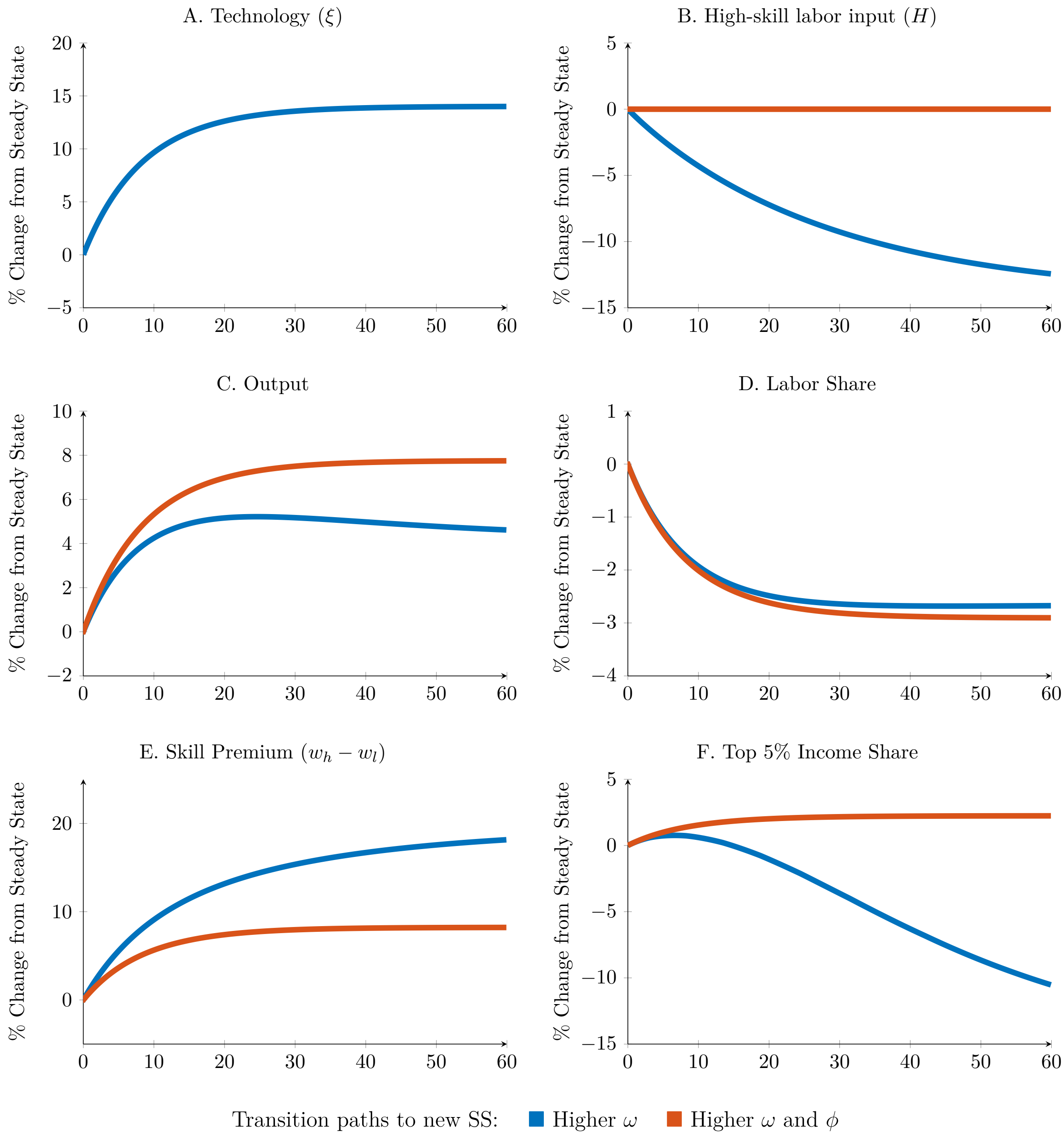
Note: This figure shows the impulse responses of key model quantities following a one-standard deviation technology shock evaluated at the steady state of the model.

Figure 6: Model: Innovation and Worker Earnings



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 7: 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 (1910–present)

$\log \text{Empl}_{i,t+h} - \log \text{Empl}_{i,t}$	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, $\eta_{i,t}$	-0.64***	-0.87***	-0.55***	-0.77***	-0.53***	-0.93***	-0.56***	-1.05***
(past decade average)	(-4.95)	(-5.27)	(-5.49)	(-5.51)	(-2.96)	(-3.95)	(-3.13)	(-4.38)
Observations	2,230	1,916	2,029	1,723	102,400	81,009	72,451	54,662
Controls								
Time FE	Y	Y	Y	Y				
Industry X Time FE					Y	Y	Y	Y
Lagged Dependent Variable			Y	Y			Y	Y

Note: Panel A of the table above reports results from regressions of the form

$$\frac{100}{h}\left(\log Y_{i,t+h} - \log Y_{i,t}\right) = \alpha_0 + \alpha_t + \beta(h)\eta_{i,t} + \rho\left(\log Y_{i,t} - \log Y_{i,t-10}\right) + \epsilon_{i,t}$$

and Panel B estimates regressions of the form

$$\frac{100}{h}\left(\log Y_{i,j,t+h} - \log Y_{i,j,t}\right) = \alpha_0 + \alpha_{j,t} + \beta(h)\eta_{i,t} + \rho\left(\log Y_{i,j,t} - \log Y_{i,j,t-10}\right) + \epsilon_{i,j,t}$$

for $h = 10, 20$ years for Census years spanning from 1910-2010. Here $Y_{i,t}$ and $Y_{i,j,t}$ are respectively the occupation i occupation-industry cell i, j share in total non-farm employment. Tech exposure $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. Standard errors are clustered by occupation in occupation-level specifications and double clustered by occupation and industry in occupation-industry level specifications; corresponding t-stats are shown atetheses. Observations are weighted by occupation or occupation-industry cell employment share at time t. Growth rates are winsorized at the 1% level.

Table 2: Comparison of technology exposure measure to machine-learning predictors

	Employment Growth (10 years)	Wage Growth (10 years)
A. Baseline		
Technology exposure, η	-1.107	-0.199
	0.203	0.047
B. Machine learning predictors of employment declines		
Mean across topics that predict employment declines	-1.140	-0.193
	0.168	0.032
First PC across topics that predict employment declines	-1.014	-0.189
	0.160	0.031

Note: The table compares the performance of our baseline measure (panel A) to the performance of a purely statistical model constructed to maximize the predictability of employment declines using patent text (panel B). To construct these predictors, 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 embedding discussed in Section 1. We then use these 500 textual factors to form a single predictor that is optimized to predict occupation declines in the MORG sample. To do so, we examine the univariate performance of each factor in predicting employment declines, and then form a linear combination (either the mean or the first principal component) of the topics that are statistically significant negative predictors at the 5% level in univariate regressions. The table reports coefficients from panel regressions of annualized wage and income growth rates over the next 10 years on these predictors, all of which are scaled to unit standard deviation. The vector of controls includes three one-year lags of dependent variable, and time fixed effects. Data come from the CPS Merged Outgoing Rotation Groups (MORG) and cover the 1985–2018 period.

Table 3: Summary Statistics: Census-CPS merged sample (worker-level data)

Variable	Mean	SD	5%	10%	25%	Median	75%	90%	95%	Observations
W2 Earnings	66,150	145,500	15,500	20,710	32,390	50,190	76,070	114,000	152,200	2,782,000
Age	40.9	7.4	29	31	35	41	47	51	53	2,782,000
Age, workers in bottom-25 income bin 25	40.0	7.6	29	30	33	40	46	51	53	632,000
Age, workers in top-5 income bin	43.2	6.8	31	33	38	44	49	52	53	109,000
Occupation-industry technology exposure (ξ)	0.647	0.971	0	0	0	0.232	0.866	2.028	2.922	1,495,000
Male	0.542	0.498	0	0	0	1	1	1	1	2,782,000
Has four-year college degree	0.344	0.475	0	0	0	0	1	1	1	2,782,000
Lifecycle-Adjusted Earnings growth, 3-years	-0.072	0.478	-0.973	-0.507	-0.129	0.008	0.129	0.312	0.472	2,773,000
Lifecycle-Adjusted Earnings growth, 5-years	-0.095	0.526	-1.116	-0.622	-0.174	0.002	0.142	0.337	0.507	2,596,000
Lifecycle-Adjusted Earnings growth, 10-years	-0.145	0.609	-1.363	-0.825	-0.280	-0.023	0.160	0.388	0.576	1,697,000

Note: The table reports summary statistics for our wage earnings data from the Census Detailed Earnings Record (DER)-CPS merged sample, which covers the 1981 to 2016 period. The sample includes all workers whose unique identifiers (PIK codes) can be matched between the DER and CPS data for CPS years between 1981 and 2016 and who satisfy labor force attachment sampling criteria. W2-Earnings are reported in terms of 2015 dollars. The occupation-industry techology exposure ξ is defined as in (10) from the main text. Patents are matched to industry of origination using information from the confindential Census SSL and LBD datasets. The variable "Has four-year college degree" denotes whether a given individual has completed a 4-year degree at the time they were observed in the CPS. Workers are required to be between the ages of 25 and 55 to be included in the sample. Lifecycle-adjusted earnings growth rates follow Guvenen et al. (2014) and are constructed following (8) and (9) in the main text. For more details on the construction of the CPS-DER matched sample and the linking of patents to industries, see appendix section A.5.

Table 4: Technology exposure and worker earnings growth

	(1)	(2)	(3)	(4)
A: All workers				
3–years earnings growth	-1.39	-1.33	-1.58	-1.72
	(0.33)	(0.29)	(0.37)	(0.37)
5–years earnings growth	-1.47	-1.34	-1.86	-1.99
	(0.38)	(0.33)	(0.39)	(0.39)
10–years earnings growth	-1.68	-1.48	-2.23	-2.35
	(0.41)	(0.36)	(0.49)	(0.44)
B: By sector				
Manufacturing (NAICS 11–33)	-2.27	-2.10	-1.88	-2.04
	0.57	0.49	0.52	0.51
Services (NAICS 42–81)	-1.00	-0.86	-1.84	-1.94
	0.47	0.37	0.48	0.47
Fixed Effects				
Industry	x	x		
Occupation	x		x	
Occupation $\times$ Year		x		x
Industry $\times$ Year			x	x

Note: This table shows the estimated slope coefficients β (times 100) from equation (11) in the main text. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. Columns (1) to (4) report coefficient estimates under different combinations of fixed effects; all specifications include prior income rank $\times$ calendar year fixed effects.

Table 5: Technology exposure and worker earnings growth, by occupation task type

	Task Intensity		High-Low
	Low	High	(p-val, %)
Manual physical	-1.21	-2.54	0.00
	0.45	0.41	
Non-routine manual and personal	-2.33	-0.80	0.00
	0.41	0.47	
Routine cognitive	-1.05	-2.51	0.00
	0.40	0.41	
Non-Routine cognitive	-2.28	-1.73	2.06
	0.41	0.42	

Note: Table shows the estimated slope coefficients β (times 100) from equation (11) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary by occupation task type. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. We report results from the specification that includes occupation $\times$ calendar year; industry $\times$ calendar year; and prior income rank $\times$ calendar year fixed effects—corresponding to column (4) in table 4. We report results across different occupation types—specifically workers in occupations that score above or below the median in terms of the Acemoglu and Autor (2011) task types.

Table 6: Technology exposure and worker earnings growth

	(1)	(2)	(3)	(4)
A: By worker education				
College	-1.47	-1.33	-1.82	-1.94
	0.39	0.33	0.39	0.38
No College	-1.54	-1.40	-1.90	-2.05
	0.41	0.37	0.43	0.42
B: By worker age				
25–35 years	-0.90	-0.78	-1.33	-1.44
	0.42	0.38	0.48	0.47
35–45 years	-1.32	-1.21	-1.73	-1.86
	0.37	0.33	0.45	0.44
45–55 years	-2.09	-1.96	-2.50	-2.63
	0.56	0.51	0.46	0.45
Fixed Effects				
Industry	x	x		
Occupation	x		x	
Occupation $\times$ Year		x		x
Industry $\times$ Year			x	x

Note: Table shows the estimated slope coefficients β (times 100) from equation (11) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary by education (panel A) or age (panel B). The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. Columns (1) to (4) report coefficient estimates under different combinations of fixed effects; all specifications include prior income rank $\times$ calendar year fixed effects, in addition to dummies for the levels of coefficient interactions.

Table 7: Technology exposure and worker earnings growth, by worker earnings rank

Worker earnings rank (rel. to occ × ind group)	All	By worker age 25–35	35–45	45–55
0–25th percentile	-1.85	-1.19	-1.53	-3.24
	0.49	0.58	0.62	0.53
25–50th percentile	-1.49	-1.09	-1.47	-1.97
	0.43	0.48	0.49	0.47
50–75th percentile	-1.85	-1.45	-1.91	-2.07
	0.39	0.47	0.44	0.45
75–95th percentile	-2.52	-2.59	-2.35	-2.70
	0.42	0.52	0.45	0.59
95–Top	-4.21	-5.20	-4.04	-4.04
	0.58	0.90	0.68	0.81
95–Top vs 0–25th percentile	-2.36	-4.02	-2.51	-0.80
(p-val, %)	0.10	0.00	0.26	30.74
95–Top vs 25–95th percentile	-2.26	-3.49	-2.13	-1.79
(p-val, %)	0.00	0.01	0.00	0.44
0–25th percentile vs 25–95th percentile	0.10	0.53	0.39	-0.99
(p-val, %)	77.70	10.85	41.10	1.31

Note: Table shows the estimated slope coefficients β (times 100) from equation (11) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary either with occupation–industry earnings rank or earnings rank interacted with age bin. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. We report results from the specification that includes occupation × calendar year; industry × calendar year; and prior income rank × calendar year fixed effects—corresponding to column (4) in table 4, also including dummies for the levels of coefficient interactions. We sort workers into income bins based on their yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th pand 95th percentile; and, 95th percentile and above.

Table 8: Technology exposure and worker earnings risk

	(1)	(2)	(3)	(4)
0–25th percentile	0.526	0.506	0.723	0.818
	0.191	0.180	0.223	0.221
25–50th percentile	0.326	0.303	0.517	0.607
	0.167	0.155	0.216	0.213
50–75th percentile	0.390	0.369	0.583	0.678
	0.172	0.159	0.205	0.205
75–95th percentile	0.618	0.601	0.809	0.906
	0.188	0.176	0.210	0.208
95–Top	1.347	1.327	1.516	1.623
	0.270	0.258	0.290	0.288
95–Top vs 25–95th percentile	0.902	0.903	0.880	0.893
(p-val, %)	0.010	0.007	0.010	0.007
95–Top vs 0–25th percentile	0.821	0.821	0.793	0.805
(p-val, %)	0.425	0.385	0.529	0.448
0–25th percentile vs 25–95th percentile	0.081	0.081	0.087	0.088
(p-val, %)	55.090	55.150	52.540	52.130
Fixed Effects				
Industry	x	x		
Occupation	x		x	
Occupation × Year		x		x
Industry × Year			x	x

Note: Table shows the estimated slope coefficients β (times 100) from a version of equation (11) in the main text, where we replace the main dependent variable—cumulative earnings growth (net of life-cycle effects) over the next 5 years—with an indicator for whether a given worker's earnings growth is beneath the 10th percentile of earnings growth for that year. Coefficients on technology exposure $\xi_{i,t}$ are allowed to vary with occupation–industry earnings rank. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. Columns (1) to (4) report coefficient estimates under different combinations of fixed effects; all specifications include prior income rank × calendar year fixed effects. We sort workers into income bins based on their yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th pand 95th percentile; and, 95th percentile and above.

Table 9: Technology exposure and worker earnings growth, heterogeneity as a function of related experience

Dependent Variable:	Earnings Growth		Tail Risk	
Related Experience Requirement:	Low	High	Low	High
0–25th percentile	-2.88	-1.16	1.28	0.50
	0.55	0.50	0.23	0.24
25–50th percentile	-1.96	-1.14	0.83	0.44
	0.50	0.45	0.24	0.23
50–75th percentile	-2.16	-1.61	0.94	0.47
	0.43	0.42	0.22	0.22
75–95th percentile	-2.22	-2.65	0.90	0.85
	0.49	0.45	0.23	0.23
95–Top	-3.71	-4.38	1.32	1.63
	0.70	0.74	0.37	0.34
95–Top vs 25–95th percentile	-1.60	-2.58	0.43	1.04
(p-val, %)	2.61	0.00	23.52	0.00
95–Top vs 0–25th percentile	-0.83	-3.22	0.04	1.13
(p-val, %)	35.35	0.01	92.46	0.04
0–25th percentile vs 25–95th percentile	-0.77	0.64	0.39	-0.09
(p-val, %)	5.09	10.22	3.00	51.52

Note: Table shows the estimated slope coefficients β (times 100) from two versions of equation (11) in the main text. In the first two columns the dependent variable is given by workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. In the second two columns the dependent variable is an indicator for whether earnings growth is below the 10th percentile for that year. Coefficients on technology exposure $\xi_{i,t}$ are allowed to vary by occupation–industry earnings rank interacted with being above or below the median on occupational related experience requirements (calculated using the O*NET related experience measure). We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. We report results from the specification that includes occupation × calendar year; industry × calendar year; and prior income rank × calendar year fixed effects—corresponding to column (4) in table 4, also including dummies for the levels of coefficient interactions. We sort workers into income bins based on their yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th pand 95th percentile; and, 95th percentile and above.

Table 10: Technology exposure and worker earnings growth: high- vs low-similarity breakthroughs

Worker earnings rank (rel. to occ × ind group)	Technology Exposure: high-similarity breakthroughs in same industry (ξ)	Low-similarity breakthroughs in same industry (ζ)
All workers	-1.53	1.05
	0.41	0.32
0–25th percentile	-1.51	1.52
	0.53	0.36
25–50th percentile	-0.96	0.77
	0.46	0.33
50–75th percentile	-1.32	0.78
	0.42	0.37
75–95th percentile	-2.08	1.12
	0.44	0.48
95–Top	-3.85	1.41
	0.59	0.48
95–Top vs 0–25th percentile	-2.34	-0.11
(p-val, %)	0.08	84.13
95–Top vs 25–95th percentile	-2.40	0.52
(p-val, %)	0.00	8.16
0–25th percentile vs 25–95th percentile	-0.05	0.63
(p-val, %)	88.45	11.19

Note: Table shows the estimated slope coefficients β (times 100) from equation (13) in the main text, where coefficients on technology exposure measures $\xi_{i,t}$ and $\zeta_{i,t}$ are allowed to vary either with occupation–industry earnings rank. The measure $\zeta_{i,t}$ is constructed to weight heavily breakthrough patents that are more textually *dissimilar* to workers' occupation task descriptions, rather than textually similar as with $\xi_{i,t}$. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. We report results from the specification that includes occupation × calendar year; industry × calendar year; and prior income rank × calendar year fixed effects—corresponding to column (4) in table 4. We sort workers into income bins based on their yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th pand 95th percentile; and, 95th percentile and above.

Table 11: 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
Human capital background risk	z	0.10
CES parameter in inner nest (technology ξ and low-skill labor L)	ρ	0.68
Share of technology in inner nest	λ	0.57
CES parameter in outer nest (high-skill labor H and ξ/L composite)	σ	-0.04
Share of high-skill labor in outer nest	μ	0.23
Size of technology improvement	κ	0.16
Annualized arrival rate of technology shocks	ω	2.06
Share of exposed workers	α	0.23
Human capital loss percentage conditional on fall	h	0.06
Annualized rate of depreciation of technology	g	0.11
Annualized likelihood of worker skill acquisition	ϕ	2.40

Note: This table reports the parameter used to calibrate the model. The first four parameters are calibrated a priori; the latter 11 parameters are chosen to fit the statistics reported in Table 12.

Table 12: Model Fit

Statistic	Data	Model
Labor share, response to ξ	-1.25	-1.16
Skill premium (p75 / p25 ratio), average	2.34	2.69
Labor productivity, response to ξ	2.81	2.67
Worker earnings growth response to ξ		
0 to 25-th percentile	-1.85	-1.13
25 to 50-th percentile	-1.49	-1.02
50 to 75-th percentile	-1.85	-2.87
75 to 95-th percentile	-2.52	-3.16
95 to 100-th percentile	-4.21	-3.59
Likelihood of large wage declines in response to ξ		
0 to 25-th percentile	0.82	1.22
25 to 50-th percentile	0.61	0.65
50 to 75-th percentile	0.68	0.65
75 to 95-th percentile	0.91	0.79
95 to 100-th percentile	1.62	1.71

Note: Table reports the fit of the model to the statistics that we target. The parameters used in our calibration are listed in Table 12.

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. (2021), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET. 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 Measuring the distance between patents and worker tasks using word embedding vectors

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

We obtain text data for measuring patent/job task similarity from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. We use the patent text data parsed from the USPTO patent search website in Kelly et al. (2021), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET. 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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[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.).

to one another and a different training instance of the same data would yield different word vectors but very similar pairwise distances between word vectors.

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

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

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 (A.6), we weigh each word vector by

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

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

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

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

The inverse-document frequency is

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

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

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:

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

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

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

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

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

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

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

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

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

A.3 Industry Innovation, Productivity, and the Labor Share

Here, we document a set of correlations regarding the joint dynamics of aggregate measures of innovation, measured productivity, and the labor share that are useful when calibrating the model. 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. We follow KPST and create an index of the degree of technological progress in industry I in year t as

$$\psi_{I,t} = \frac{1}{N_t} \sum_{p \in B_t} \alpha_{I,p}. \tag{A.13}$$

To determine the set of breakthrough patents that are relevant to a given industry, KPST map patents to industries based on their CPC technology class using the probabilistic mapping constructed by Goldschlag et al. (2020). Here, $\alpha_{j,p}$ denotes the probability of breakthrough patent p being assigned to industry I. Last, following KPST and given the relatively long time series considered we scale (A.13) by US population N_t so that $\psi_{I,t}$ is relatively stationary over time, and we normalize to unit standard deviation.

Using the KPST index of technological breakthroughs, we explore the extent to which they are related to future labor productivity and the labor share. In particular, we estimate the following specification

$$\log X_{I,t+h} - \log X_{I,t} = \alpha_t + \beta\, \psi_{I,t} + c\, \mathbf{Z}_{I,t} + \varepsilon_{I,t}, \qquad k = 1 \ldots T \text{ years.} \tag{A.14}$$

We focus on four outcome variables: output (specifically, value-added); employment; labor productivity (value-added per worker); and the labor share. We examine horizons of up to $T = 6$ years. In addition to calendar year fixed effects α_t, we include in the vector of controls $\mathbf{Z}$ the lagged 5-year growth rate of the outcome variable.

Appendix Figure A.3 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 the KPST 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.

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. We also 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.

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 3 years, thus effectively dropping observations where the CPS interview date is older than 3 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 firm's 4-digit NAICS code. In particular, we use two SSEL patent–assignee crosswalks: the newer Business Dynamics Statistics of Patenting Firms database (BDS-PF) and an older patent-SSL crosswalk created by Kerr and Fu (2008). The BDS-PF links are available starting with the 2000 SSL. We use the BDS-PF firmid-patent links for any patents for which it is available. Otherwise we use the union of links created by Kerr and Fu (2008) from the 1976–1999 SSL data. We obtain NAICS codes by using the "firmid" identifier to join with the Longitudinal Business Database (LBD), and we use the 2012 version of Fort and Klimek (2018) NAICS codes, which are adjusted to improve industry comparability over time. In cases where a firmid matches to multiple NAICS codes we apply the 4-digit NAICS code of highest employment based off the LBD.

To allow the effects to vary with prior income, we assign workers into five groups based on their income in the previous year compared to workers in the same occupation and NAICS4 industry. These groups are defined based on the following percentiles of prior income [0%, 25%), [25%, 50%), [50%, 75%), [75%, 95%), [95%, 100%] calculated within industry–occupation cells. In the (uncommon) case when NAICS4 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. (2021) as described in section 1), 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. Appendix figure A.4 shows some example word clouds for a selection of four of the top 500 patent topics.

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
$$\psi_{i,k,t} = \frac{O_{k,i}}{N_t} \sum_{j \in \widehat{\Gamma}_t} P_{k,j} \tag{A.15}$$

As before we only sum over breakthrough patents and normalize by U.S. population in year t (denoted by N_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.15 looks a bit like our construction of $\eta_{i,t}$ in equation 3, 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 4. 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 1. 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%.

Table 2 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.

In terms of magnitudes, the employment and wage declines predicted by this statistical displacement factor are comparable to our baseline measure—that is, 1.11% vs 1.14% employment declines at the 10-year horizon and 0.20% vs 0.19% decline in wage earnings for the version of the statistical displacement factor that takes the mean across topics. The correlation between our baseline measure $\eta_{i,t}$ and the statistical predictor constructed to represent exposure to labor-saving technologies is also quite high, at approximately 73 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

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

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

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

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

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

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. $m_{i,t}$ is an i.i.d. random variable which governs whether a worker learns or forgets conditional on a dM shock; skills grow by m% with probability $1-z$ and shrink by m% with probability z. $d_{i,t}$ 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 in the interior of the θ interval, we have

$$E(\theta_{i,t}) = \theta_{i,t-1} + m(1-2z)\,\phi\,\theta_{i,t-1} - \alpha\,h\,\theta_{i,t-1}$$

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.

$$E(\theta_{i,t}|s=0) = \underbrace{(1-\delta)}_{\substack{\text{adj for}\\ \text{worker exit}}}\,[\theta_{i,t-1} - \underbrace{\phi\,\theta_{i,t-1}}_{\substack{\text{outflow from}\\ \text{learning/obsolescence}}}] + \underbrace{\phi(1-z)(1-\delta)\,\theta_{i-1,t-1}}_{\substack{\text{inflow from}\\ \text{learning}}} + \underbrace{\phi\,z\,(1-\delta)\theta_{i+1,t-1}}_{\substack{\text{inflow from}\\ \text{skill obsolescence}}}$$

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

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

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 $\phi(1-z)$ to a higher gridpoint as workers acquire new skills, and gain shares $\phi(1-z)$ and $\phi\,z$ from the gridpoints below and above, respectively.

We can represent this as a Markov-switching 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

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

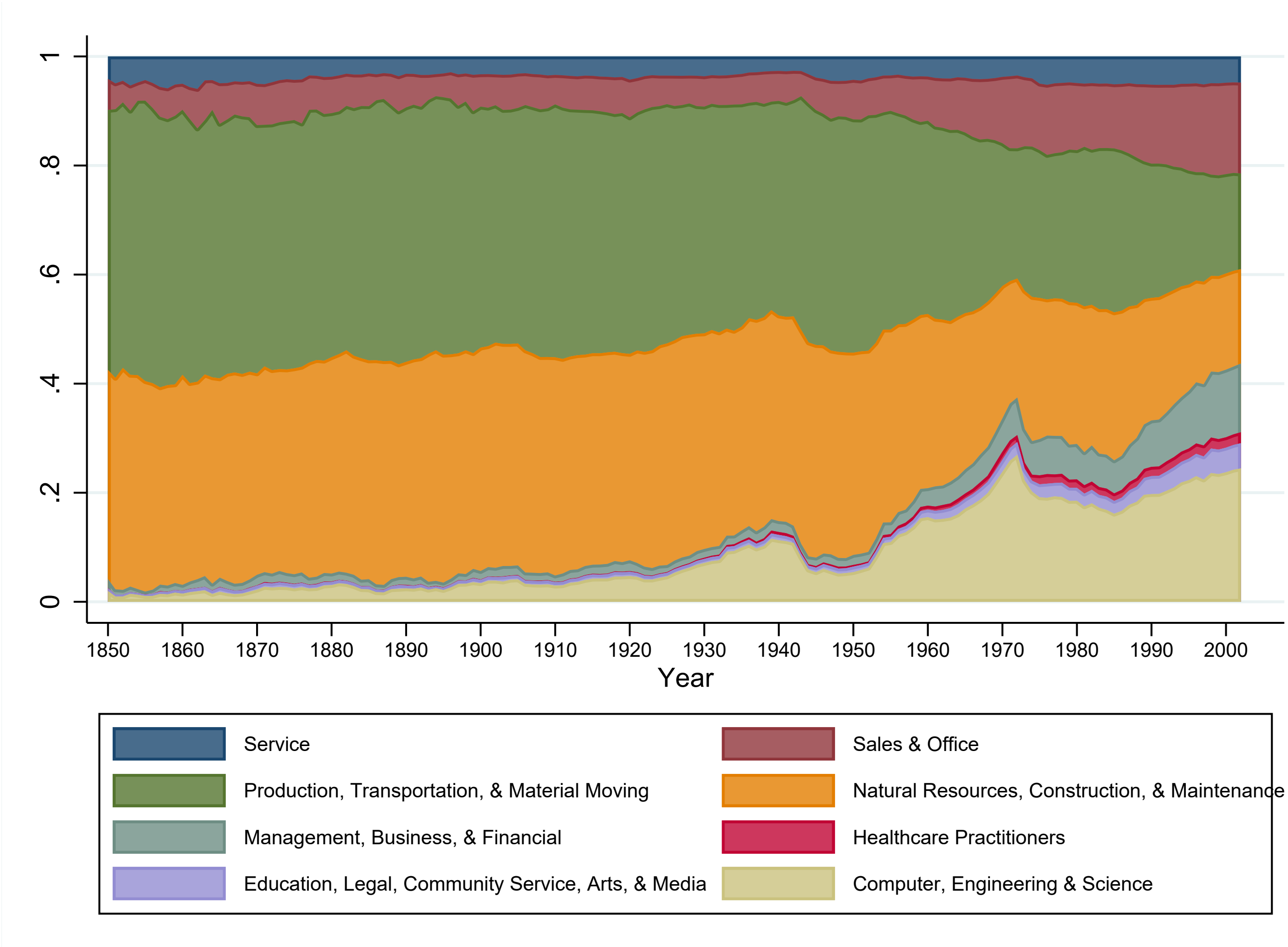
and

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

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.

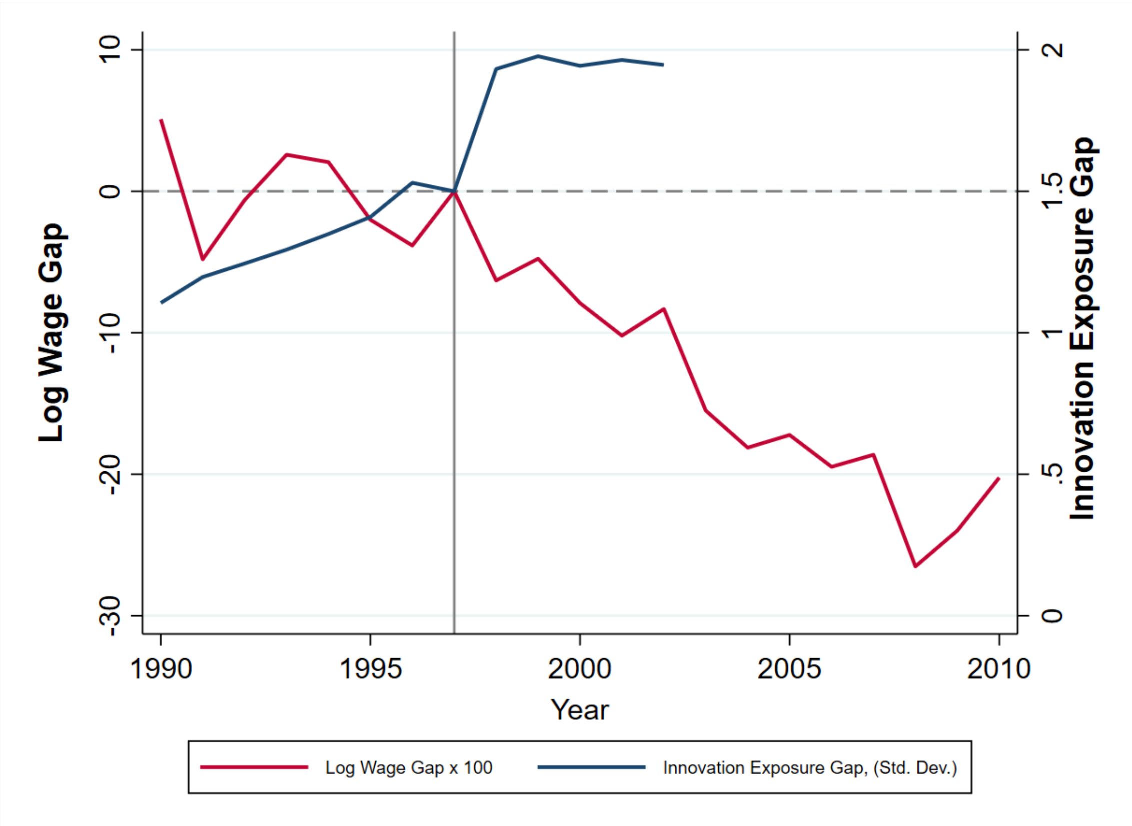
Appendix Figures and Tables

Figure A.1: Technological Exposure, composition by major occupation group



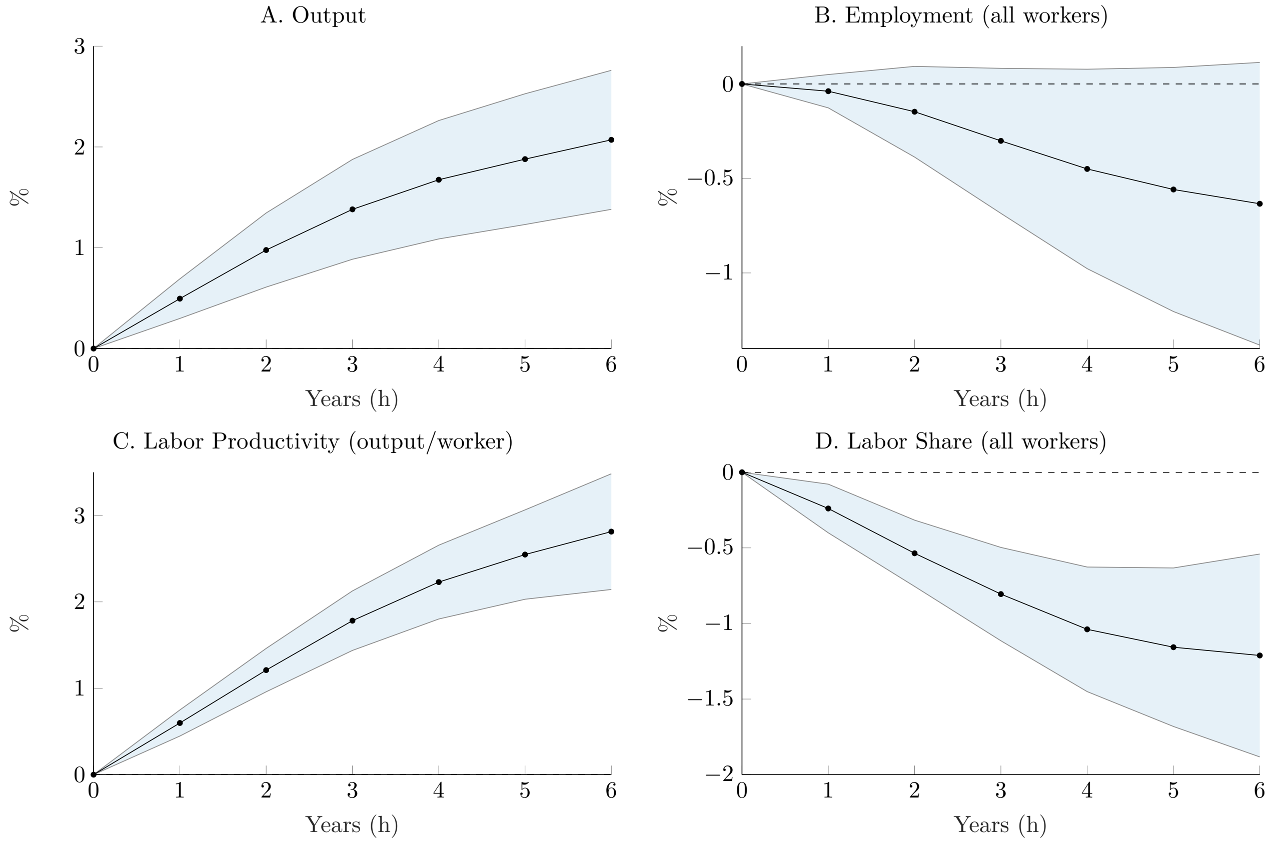
Note: This figure plots the average of our occupation-level innovation exposure index, $\eta_{i,t}$, where $\eta_{i,t}$ has been averaged separately within eight broad occupation groups. The occupation group averages are re-scaled each year so that the total across all groups sums to one in the given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

Figure A.2: Example: Order Clerks versus All Other Clerk Occupations



Note: This figure plots the gap between average log wages ($\times$ 100) for order clerks versus all other clerks in (left axis) and the gap between our innovation exposure measure $\eta_{i,t}$ for order clerks versus all other clerk occupations. The log wage gap is normalized to 0 in 1997. The wage data come from the CPS Merged Outgoing Rotation Groups.

Figure A.3: 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 A.4: Sample Patent Topic Word Clusters

The above are four of the top 500 most important textual topics for patents resulting from the LSH approximate nearest neighbors routine used to separate words into clusters as described in appendix section section A.6 and section 2.3 in the main text. The relative size of the word corresponds to the importance of that word within the topic.

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 25 Occupations by Average $\eta_{i,t}$	Bottom 25 Occupations by Average $\eta_{i,t}$
Production checkers, graders, and sorters in manufacturing	Funeral directors
Miscellaneous precision workers, n.e.c.	Dancers
Punching and stamping press operatives	Barbers
Machinery maintenance occupations	Sheriffs, bailiffs, correctional institution officers
Rollers, roll hands, and finishers of metal	Pest control occupations
Production helpers	Optometrists
Lathe and turning machine operatives	Actuaries
Typesetters and compositors	Podiatrists
Metal platers	Bartenders
Extruding and forming machine operators	Lawyers and judges
Grinding, abrading, buffing, and polishing workers	Bakers
Programmers of numerically controlled machine tools	Clergy
Machine feeders and offbearers	Plasterers
Production supervisors or foremen	Registered nurses
Laborers, freight, stock, and material handlers, n.e.c.	Shoemaking machine operators
Nail, tacking, shaping and joining mach ops (wood)	Dental Assistants
Drilling and boring machine operators	Musicians and composers
Sheet metal workers	Butchers and meat cutters
Millwrights	Garbage and recyclable material collectors
Packers, fillers, and wrappers	Food preparation workers
Cementing and gluing machne operators	Physicians
Weighers, measurers, and checkers	Hotel clerks
Photographic process machine operators	Atmospheric and space scientists
Forge and hammer operators	Pharmacists
Drillers of earth	Subject instructors, college

Note: This table ranks occupations (defined at the David Dorn revised Census occ1990 level) by their time series average technology exposure score $\eta_{i,t}$ from (3) in the main text. The first column gives the top 25 most exposed occupations on average, while the second 25 columns give the top 25 least exposed. The sample period spans 1850–2002.

Table A.4: Technology and Labor Market Outcomes: Comparison to Other Measures

	A. Employment						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tech Exposure 1980 Rank	-4.48***	-2.79***				-3.78***	-2.23***
	(-10.30)	(-4.82)				(-6.03)	(-3.11)
Routine Task Intensity Rank			-2.16***			-2.31***	-1.64***
			(-4.88)			(-3.82)	(-3.05)
Robot Exposure Rank				-1.69***		0.52	-0.15
				(-3.95)		(0.73)	(-0.24)
Software Exposure Rank					-1.20**	-0.34	0.050
					(-2.28)	(-0.47)	(0.07)
Industry Fixed Effects		X	X	X	X		X
Observations	17154	17154	17154	17154	17154	17154	17154
R^2	0.17	0.43	0.43	0.42	0.41	0.20	0.45
	B. Wages						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tech Exposure 1980 Rank	-5.28***	-3.24***				-4.41***	-2.47***
	(-11.74)	(-5.66)				(-6.96)	(-3.40)
Routine Task Intensity Rank			-2.60***			-2.48***	-1.85***
			(-6.33)			(-4.59)	(-4.00)
Robot Exposure Rank				-2.27***		-0.13	-0.70
				(-5.08)		(-0.19)	(-1.25)
Software Exposure Rank					-1.43***	0.057	0.32
					(-2.64)	(0.08)	(0.50)
Industry Fixed Effects		X	X	X	X		X
Observations	17154	17154	17154	17154	17154	17154	17154
R^2	0.21	0.47	0.47	0.46	0.44	0.25	0.50

Note: This table shows results from estimating

$$\frac{100}{h}\Big(\log Y_{i,j,t+h} - \log Y_{i,j,t}\Big) = \alpha + \alpha_j + \beta \eta^{\text{Pctile}}_{i,1980} + \delta X^{\text{Pctile}}_i + \epsilon_{i,j}$$

Here i indexes occupations and j indexes industries; we report results for $h = 32$ 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). We additionally include the routine-task intensity from Acemoglu and Autor (2011) and the measure of exposure to robots or software from Webb (2019), depending on the specification. Because the Webb (2019) exposure measures are reported in percentile ranks, we convert our index of technological exposure in 1980 to its corresponding cross-sectional percentile rank, and also do the same for the routine-task intensity. Coefficients are multiplied by 100 for readability, and observations are weighted by employment share in 1980.

Table A.5: 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.6: Technology exposure and worker earnings growth by income rank, alternative income rankings

	Income Ranking Method			
	Baseline	Firm-Adjusted	Residual Earnings	
	(1)	(2)	(3)	(4)
0–25th percentile	-1.85	-2.09	-2.04	-1.82
	0.49	0.49	0.44	0.61
25-50th percentile	-1.49	-1.71	-1.41	-1.45
	0.43	0.42	0.38	0.49
50-75th percentile	-1.85	-1.80	-1.81	-1.46
	0.39	0.40	0.37	0.55
75-95th percentile	-2.52	-2.43	-2.12	-2.01
	0.42	0.41	0.39	0.60
95–Top	-4.21	-3.55	-3.63	-3.89
	0.58	0.56	0.65	0.86
95–Top vs 25–95th percentile	-2.26	-1.46	-1.85	-2.25
(p-val,%)	0.00	0.33	0.02	0.04
95–Top vs 0–25th percentile	-2.36	-1.57	-1.59	-2.08
(p-val,%)	0.10	0.01	1.08	0.55
0-25th percentile vs 25–75th percentile	0.10	-0.11	-0.27	-0.18
(p-val,%)	77.70	72.26	40.20	66.05

Note: This table shows the estimated slope coefficients β (times 100) from equation (11) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary with worker earnings rank. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. Columns (1) through (4) use different methods for ranking workers based on earnings. Column (1) represents our baseline method of sorting workers each year within occupation–industry. Column (2) adjusts workers' earnings for average firm wages by subtracting off the log average wage of a worker's employer in the Longitudinal Business Database from the worker's log wage. Column (3) residualizes log earnings with respect to yearly fixed effects for occupation–industry; commuting zone; and 10-year age bin (25-35, 35-45, 45-55) interacted with gender. Column (4) restricts to the roughly one-fifth of workers in our sample who are in-universe for the CPS union membership question, and residualizes log earnings with respect to yearly fixed effects for occupation, industry, worker union status, and commuting zone. All specifications include industry × year, occupation × year, and within occupation–industry income bin × year fixed effects, in addition to dummies for the levels of coefficient interactions.

Table A.7: Technology exposure and worker earnings growth by income rank, comparison across unionized vs non-unionized workers

	Industry Unionization		Worker In Union	
Worker earnings rank (rel. to occ × ind group)	Low	High	Non-Union	Union
0–25th percentile	-2.18	-1.88	-2.29	-1.00
	0.63	0.63	0.81	1.69
25–50th percentile	-1.92	-1.10	-1.25	-2.87
	0.52	0.52	0.70	1.00
50-75th percentile	-2.26	-1.55	-2.24	-2.56
	0.52	0.47	0.68	0.95
75-95th percentile	-2.85	-2.21	-2.63	-3.67
	0.55	0.59	0.68	1.35
95-Top	-5.50	-3.20	-4.84	-4.01
	0.78	0.89	0.97	1.74
95–Top vs 25–95th percentile	-3.16	-1.62	-2.80	-0.98
(p-val,%)	0.00	7.03	0.00	38.87
95–Top vs 0–25th percentile	-3.32	-1.32	-2.55	-3.01
(p-val,%)	0.01	24.19	0.13	6.69
0–25th percentile vs 25–95th percentile	0.16	-0.26	-0.25	2.03
(p-val,%)	68.17	69.29	50.63	16.02

Note: This table shows the estimated slope coefficients β (times 100) from equation (11) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary with occupation–industry earnings rank and either industry unionization rates or worker union status. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. In the first two columns report results we sort workers into yearly within occupation–industry income bins, interacted with whether the average reported union membership status by workers in their industry is above or below the median. In the last two columns we restrict the sample to the roughly one-fifth of workers in our sample who are in-universe for the CPS union membership question, and interact within workers' occupation–industry income bins with individual union status. All specifications include industry × year, occupation × year, and within occupation–industry income bin × year fixed effects, in addition to dummies for the levels of coefficient interactions.

Table A.8: Technology exposure and worker earnings risk

Worker earnings rank (rel. to occ $\times$ ind group)	Earnings growth $\leq$ p10	ExitSR: 1 year of zero of zero W2 income	ExitLR: 3 consecutive years of zero W2 income
0–25th percentile	0.818	0.771	0.330
	0.221	0.208	0.090
25–50th percentile	0.607	0.425	0.180
	0.213	0.183	0.088
50–75th percentile	0.678	0.513	0.272
	0.205	0.184	0.095
75–95th percentile	0.906	0.577	0.272
	0.208	0.196	0.094
95–Top	1.623	0.679	0.297
	0.288	0.278	0.128
95–Top vs 25–95th percentile	0.893	0.174	0.055
(p-val, %)	0.007	37.610	47.670
95–Top vs 0–25th percentile	0.805	-0.092	-0.034
(p-val, %)	0.448	76.190	76.610
0–25th percentile vs 25–95th percentile	0.088	0.266	0.089
(p-val, %)	52.130	9.656	14.320

Note: Table shows the estimated slope coefficients β (times 100) from a version of equation (11) in the main text, where we replace the main dependent variable—cumulative earnings growth (net of life-cycle effects) over the next 5 years—with an indicator for whether a given worker's earnings growth is beneath the 10th percentile of earnings growth for that year (first column); an indicator for whether a given worker reports no W2 earnings in any year in the next 5 years (second column); and, an indicator for whether a given worker reports no W2 earnings for any consecutive 3-year period within the next 5 years (third column). Coefficients on technology exposure $\xi_{i,t}$ are allowed to vary with occupation–industry earnings rank. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. All specifications include industry $\times$ year, occupation $\times$ year, and within occupation–industry income bin $\times$ year fixed effects, in addition to dummies for the levels of coefficient interactions.

Table A.9: Technology exposure and worker earnings growth, conditioning on continuing employment

Worker earnings rank (rel. to occ × ind group)	All Workers	Continual Employment	
		ExitSR=0 (no years with missing W2)	ExitLR=0 (no 3-cons years with missing W2s)
0–25th percentile	-1.848	-1.201	-1.493
	0.494	0.371	0.454
25–50th percentile	-1.493	-1.216	-1.360
	0.433	0.330	0.408
50–75th percentile	-1.847	-1.352	-1.568
	0.393	0.290	0.352
75–95th percentile	-2.516	-1.960	-2.262
	0.425	0.320	0.386
95–Top	-4.212	-3.549	-3.938
	0.585	0.465	0.514
95–Top vs 25–95th percentile	-2.260	-2.040	-2.208
	0.000	0.000	0.000
95–Top vs 0–25th percentile	-2.364	-2.348	-2.445
	0.101	0.001	0.008
0–25th percentile vs 25–95th percentile	0.104	0.308	0.237
	77.700	22.730	42.540

Note: This table shows the estimated slope coefficients β (times 100) from equation (11) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary with worker earnings rank. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 5 years. We report standard errors clustered at the industry (NAICS 4-digit) level beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. The first column replicates income sorts for our baseline regression sample, as in column (4) of Table 7. The second column removes workers reporting no W2 earnings in any year in the next 5 years, and the third column removes workers who report no W2 earnings for any consecutive 3-year period within the next 5 years. All specifications include industry × year, occupation × year, and within occupation–industry income bin × year fixed effects, in addition to dummies for the levels of coefficient interactions. We sort workers into income bins based on their yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th pand 95th percentile; and, 95th percentile and above.