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: August 2022

Abstract

We develop a measure of workers' technology exposure that relies only on textual descriptions of patent documents and the tasks performed by workers in an occupation. Our measure appears to identify a combination of labor-saving innovations but also technologies that may require skills that incumbent workers lack. Using a panel of administrative data, we examine how subsequent worker earnings relate to workers' technology exposure. We find that workers at both the bottom but also the top of the earnings distribution are displaced. Our interpretation is that low-paid workers are displaced as their tasks are automated while the highest-paid workers face lower earnings growth as some of their skills become obsolete. 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, Anders Humlum, Michael Peters, Pascual Restrepo, Jonathan Rothbaum, Miao Ben Zhang, and seminar participants at University of Amsterdam, Boston University, Columbia GSB, FIRS, Johns Hopkins, HKUST, Labor and Finance Group, NBER (EFG, PRMP), Macro-Finance Society, MIT Sloan, Michigan State, the Econometric Society, Rice University, University of Rochester, the Society of Economic Dynamics, University College London, University of Illinois at Urbana Champaign, University of Toronto, Tsinghua PBC, WFA, and Wharton for valuable discussions and feedback. We thank Carter Braxton, Will Cong, and Jonathan Rothbaum for generously sharing code. 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 prospect of technological displacement of labor.[1] New technologies can displace incumbent workers either because certain tasks can now be performed instead by a machine or software (labor-saving technologies), or because these technologies require new skills that incumbent workers lack. These concerns have been exacerbated by recent breakthroughs in new technologies which have occurred contemporaneously with an increase in income inequality and a fall in the labor share of aggregate output. Yet, despite the importance of these issues, systematic evidence for technological displacement remains elusive as measurement lags theory.

We propose a new measure of workers' exposure to technological innovation and examine its relation with individual worker outcomes. We identify workers' technology exposure based on the similarity between the textual description of the tasks performed by an occupation and that of major technological breakthroughs. We identify the latter using the methodology of Kelly, Papanikolaou, Seru, and Taddy (2021) who define a breakthrough innovation as one that is both novel (i.e. distinct from prior patents) and impactful (i.e. related to subsequent patents). We then estimate the similarity between a breakthrough innovation and workers' task descriptions using natural language processing. By exploiting the timing of patent grants, we measure the extent to which individual workers are exposed to major technological breakthroughs at a given point in time. Though much of the literature has focused on the effect of automation on low-skill workers, we provide new evidence consistent with the displacement of higher-paid workers due to skill obsolescence—a pervasive phenomenon that applies to a broader set of occupations than those emphasizing routine or easily automated tasks.

Prior to examining the link between our exposure measure and individual outcomes, we validate that our methodology can reliably identify labor-displacive technologies. First, we analyze specific examples. It appears that our measure picks up two types of technologies: labor-saving/automation innovations (i.e., those which explicitly replace labor with capital); and new vintages of technologies that, while potentially complementary to labor, may require skills that incumbent workers lack.[2] Second, we examine how our measure correlates with labor market outcomes at the occupation level using public-use Census micro-data that are available over longer horizons. When we compare

[1] 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 (the Luddites) worried that mechanizing manufacturing (and the unskilled laborers operating the new looms) would rob them of their means of income. In 1930, Keynes worried about technological unemployment: "we are being afflicted with a new disease of technological unemployment...due to our discovery of means of economising the use of 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 new technologies by the year 2030.

[2] For example, our technology exposure for glass blowers takes a relatively high value in the early 1900s because of similarity with breakthrough patents such as US patent number 814,612, titled "method of making glass sheets." This patent relates to a technology for making glass (the cylinder machine) which allowed glass manufacturers to replace the labor of skilled hand glass blowers in favor of a highly mechanized production process that now required skilled machine operators instead. Jerome (1934) documents that, by 1905 many hand plants had gone out of business and the wages of blowers and gatherers were reduced by 40 per cent.

workers in the same industry across occupations differentially exposed to technology improvements, we see that an acceleration in the rate of breakthrough innovations is associated with future declines in employment and average wage earnings. The fact that both wages and employment decline subsequent to an increase in technology exposure suggests that the directed development of labor-saving technologies towards occupations expected to have scarcity of workers is unlikely to be the main driver of the time-series variation in our measure. Third, our technology exposure measure performs about as well as a statistical predictor explicitly constructed to maximize the in-sample predictability of employment declines.

Examining the time series variation in our measure, we see that from the 1850s to the 1980s, the occupations most exposed to technology are those that emphasize manual physical tasks (that is, blue collar workers). By contrast, the innovations of the late 20th and early 21st century have become relatively more related to cognitive tasks—thus, occupations emphasizing routine cognitive tasks (that is, white collar workers) have been increasingly more exposed. This pattern is partly driven by the increased prevalence of breakthrough patents related to the ICT revolution. Last, occupations that are associated with interpersonal tasks have consistently low exposures to innovation throughout the entire sample period—which further reinforces the importance of social skills in the labor market documented in Deming (2017).

Next, we examine how our measure of labor-displacive technology exposure relates to individual worker outcomes. To do so, we rely on employer-employee matched administrative earnings records from the US Social Security Administration, starting in the early 1980s, which are linked with information on their 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 introduction of related technologies, then estimate how her earnings evolve in future years even if she switches employers, industries, and/or occupations. Our empirical analysis leverages the granularity of our patent-occupation measures to exploit variation at the industry-occupation level driven by relative differences in the rate at which firms in different industries develop new technologies related to a given occupation at a point in time.

We find that, in response to a standard deviation increase in technology related to her occupation, 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. Perhaps surprisingly, it is also 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 wage earnings responses across age and income levels. In particular, older workers are significantly more affected than younger workers. In addition, workers at the top of the earnings distribution—relative to their peers in the same occupation and in the same industry—experience a significantly greater decline in earnings relative to the average worker (over 0.04 log points). This income pattern persists when we rank workers based on their income relative to the firm's average wage, and thus is unlikely to be driven by between-firm differences; or when we rank workers based on their residual income net of characteristics that likely capture variation in wages unrelated to worker skill, such as age, gender, work location, or unionization status.

The income patterns we uncover vary significantly across job types. Specifically, we see that the response of wage earnings of low-income workers to our exposure measure is significantly larger in occupations that emphasize manual and routine cognitive tasks—which suggests task automation is a relevant risk for these workers. Consistent with this view, job separations account for most of the earnings losses for these lower-paid workers. By contrast, we find only minor differences in the response of the earnings of the highest-paid workers across these task categories. This pattern, together with the fact that job separations account for only a modest fraction of the earnings losses for the top workers suggests that forces beyond task automation are at play. Our interpretation is that technology can lead to some, but not all, of the workers' skills becoming obsolete. Consistent with this view, we find that the income pattern we identify is significantly stronger in occupations that require a greater amount of related experience (a proxy for skill specificity).

In sum, our preferred interpretation of these patterns is that our measure picks up a combination of labor-saving innovations and also technologies that could be complementary to labor but may require skills that incumbent workers lack. Workers at the top of the earnings distribution are displaced because part of their accumulated skills (human capital) is specific to a particular technology vintage (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). As technology improves, some of these skills can become obsolete in the new vintage.[3] In Violante (2002) for instance, the amount of skill loss is greater if the new technology is meaningfully different from the old one. Given our focus on breakthrough patents—which by construction are relatively distinct (novel) from prior technologies—this mechanism is particularly relevant. As a result, workers who have accumulated the most skill in the existing technology therefore have the most to lose when new breakthroughs arrive.

We formalize this intuition in a model in which improvements in technology are complementary to high-skill labor but act as a substitute for low-skill labor (similar to Krusell, Ohanian, Ríos-Rull, and Violante, 2000). We extend the model to allow for human capital to have a vintage-specific

[3]Skill obsolescence can be a stand-in for several economic mechanisms. One possibility is that formerly specialized tasks are standardized leading to prior expertise being less valuable (as in Acemoglu, Gancia, and Zilibotti, 2012). Another possibility is that a different set of skills is required. In this case, even if a new technology complements the skills of some workers, existing incumbents may be unable to effectively utilize it.

component: 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. Top workers – who receive a larger fraction of income as compensation for specific skills – may experience lower average earnings growth following periods of technological advances if the increase in the likelihood of displacement is sufficiently high. Low-skill workers are displaced both because technology is a substitute for their skills but also due to increased competition from (formerly) high-skill workers who supply low-skill labor after being displaced. Overall, high-skill labor becomes more scarce and low-skill labor becomes more abundant, increasing the skill premium and (potentially) income inequality. The model quantitatively replicates the facts we document in the data.

The calibrated model allows us to study the implications of an acceleration of the rate of innovation in the economy. 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 this case, more displacement leads to a decline in the aggregate supply of high-skill labor. In the second case, we also increase the rate at which workers acquire new skills so that the overall level of skilled human capital stays constant. In both scenarios, the shift in technology increases output, lowers the labor share, and increases 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 case where only technology improves, 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 case where both technology and the rate of skill acquisition increase, the skill premium increases by significantly less, while income inequality increases in both the short and long run as skills are less likely to be displaced. In both cases the behavior of the skill premium is not a sufficient summary statistic for income inequality; technology moves both skill prices but also the effective quantity of skills due to displacement.

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, Levy, and Murnane (2003) document the secular decline in occupations specializing in routine tasks—which are easier to automate—starting in the late 20th century.[4] Consistent with the prevailing view on job polarization (Autor and Dorn, 2013), we find that occupations at the middle of the skill (income) distribution have been significantly more exposed to technology in the post-1980 period than workers at either the top or the bottom of the distribution. Webb (2020) also analyzes the similarity between patents identified as being related to robots, AI, or software and occupation task descriptions. and relates these exposures to

[4] Graetz and Michaels (2018); Dauth, Findeisen, Suedekum, and Woessner (2021); Koch, Manuylov, and Smolka (2021); Aghion, Antonin, Bunel, and Jaravel (2021); Bessen, Goos, Salomons, and van den Berge (2022); Acemoglu and Restrepo (2020, 2018, 2021) are additional examples of work on the causes and effects of adopting robots and other automation technologies.

cross-sectional differences in employment growth across occupations.[5] These papers primarily focus on group-level worker outcomes. By contrast, Akerman, Gaarder, and Mogstad (2015) and Humlum (2019) provide in depth analyses of impacts of adoption of broadband internet and industrial robots on incumbent workers' subsequent outcomes, respectively, leveraging employer-employee matched data from Scandinavia.

Our work complements this literature along several dimensions. First, our focus on individual workers allows us to study sources of ex-ante heterogeneity. This helps expose the economic forces that may drive worker displacement beyond task automation, which has been the primary focus of this literature. Specifically, our finding that high-skill workers are exposed to the risk of technology rendering their expertise obsolete occurs across a broad set of occupations and industries—beyond those emphasizing routine tasks. Our work thus complements existing studies which emphasize displacement. Goldin and Katz (2008) provide examples of how task standardization can displace incumbent skilled workers. Atack, Margo, and Rhode (2019, 2022) analyze how workers' tasks transitioned from hand to machine production in the late 19th century. More recently, Feigenbaum and Gross (2020) focus on the particular case of telephone operators and show that incumbent workers were more likely to be in lower-paying occupations following the adoption of mechanical switching technology by AT&T. We contribute to this literature by providing a measure of occupational exposure to technical change that is broad both in terms of technologies and the time period it considers (as early as the middle of the 19th century).

Our technology exposure measure is constructed largely from the perspective of incumbent workers and is primarily intended to capture technological displacement of existing tasks. Building on our work, Autor, Chin, Salomons, and Seegmiller (2022) examines the extent to which technology facilitates the creation of new tasks and occupations. Specifically, they construct a measure of labor-augmenting technologies by calculating the overlap between patent texts and the micro-titles from the Census Alphabetical Index associated with each industry and occupation. They find that augmentation innovations predict employment expansions and the proliferation of new tasks within exposed occupations.

More broadly, 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 et al., 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, 2021; Caunedo, Jaume, and Keller, 2021; Hemous and Olsen, 2021). The main focus of these papers is how

[5] In related work, Mann and Püttmann (2018); 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.

technology affects differences in wages between groups with different ex-ante skill levels (typically education). The facts we document are not at odds with the view that recent technological advances are skill-biased. In fact, when we focus on breakthrough technologies in the worker's industry that have *low levels of similarity* to the worker's tasks we find that earnings of all workers subsequently increase—with workers at the top experiencing larger increases than the average worker. Our contribution is to show that some breakthrough technologies (those closely related to worker's tasks) are instead labor-displacive even for highly-paid workers performing non-repetitive tasks. These technologies are still complementary to skill, however the skills required may not be the same as the skills that incumbent workers have. Unlike much of this literature, we do not consider worker skill to be a fixed characteristic of the worker: 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 the literature on vintage specificity of human capital (Chari and Hopenhayn, 1991; Violante, 2002; Deming and Noray, 2020; Hombert and Matray, 2021) 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).

Last, our work has important implications for the drivers of labor income risk. Though risk is not the primary focus of our study, we reach a similar conclusion as Kogan, Papanikolaou, Schmidt, and Song (2020): the highest-paid workers face considerably greater risk in their labor income as a result of technological innovation than the average worker. Though this conclusion is similar, these two papers ask different questions. 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; thus, they examine the dynamics of wage earnings in response to innovation by the workers' own firm or its competitors in the product market. 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; appendix sections A.1 and A.2 contain further details.

1.1 Data

Our raw text data comes from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. Given that the main empirical analysis in this paper focuses on worker-level outcomes in the post-1980 sample, we use the task descriptions from the 1991 DOT—which is mostly identical to the 1977 DOT version beyond the addition of some IT-related occupations—rather than more recent versions from O*NET. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we use a crosswalk from DOT occupations to the considerably coarser and yet still detailed set of 6-digit Standard Occupation Classification (SOC) codes from O*NET. We then combine all tasks for a given occupation at the 2010 SOC 6-digit level into one occupation-level corpus. We 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. See Appendix A.1 for a step-by-step description of how we clean and prepare the patent texts and occupation task descriptions for numerical analysis.

1.2 Similarity between occupation tasks and patents

Our objective is to measure textual similarity between occupation tasks and technologies as described in patents; however, a complication arises from the fact that the language 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 words, and hence allow for synonyms or other relatedness in word meanings. Specifically, we use word embeddings—also termed word vectors, since they represent word meanings as dense vectors. Word embeddings are constructed so that the distance between two word vectors is directly related to the likelihood these words capture a similar meaning. We use the word vectors provided by Pennington, Socher, and Manning (2014), which contains a vocabulary of 1.9 million word meanings. We briefly describe the approach here; Appendix A.2 contains a detailed exposition.

We represent each document (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:

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

We choose the weights $w_{i,k}$ to emphasize important words in the document. We follow the literature on natural language processing and 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. Last, we calculate the cosine similarity to measure the similarity between patent i and occupation j,

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

We perform two adjustments to the raw measure of similarity (2). First, we remove yearly fixed effects. We do so in order to account for language and structural differences in patent documents over time; patents have become much longer and use much more technical language over the sample period. Second, we impose sparsity: 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.

In sum, we use a combination of word embeddings and TF-IDF weights in constructing a distance metric between a patent document (which includes the abstract, claims, and the detailed description of the patented invention) and the detailed description of the tasks performed by occupations. Our methodology is conceptually related, though distinct, to the method proposed by Webb (2020), who also analyzes the similarity between patents and job tasks.[6]

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, in Section 2.1, we examine specific examples of breakthrough technologies and identify the most related occupations. We see that our measure primarily picks up two types of technologies: a) labor-saving innovations and b) technologies that could be complementary to labor, but may require skills that incumbent workers lack. Second, in Section 2.2, 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.

Overall, it appears that our measure identifies primarily labor-displacive innovations; that is innovations in which tasks of certain workers are replaced, either by the technology itself (labor-saving technologies) or by shifting labor demand towards new types of workers whose skills may complement the new technology. In both cases, incumbent workers can be displaced by the technology, hence we use the term labor-displacive. That said, one remaining question is whether the way we construct our technology exposure measure is indeed well-suited to identifying technologies that displace incumbent workers. In Section 2.3, we present evidence based on a purely statistical prediction exercise that such counteracting effects are likely to be small, and conclude that our measure primarily captures labor-displacive innovations. Last, Section 2.4 compares to existing (cross-sectional) measures of technology exposure, while Section 2.5 documents how workers' technology exposure varies across occupations and time.

2.1 Examples

A key advantage of our measure is that it allows us to study very different technologies across long periods of time. We consider 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

[6] Webb (2020) focuses on similarity in verb-object pairs in the title and the abstract of patents with verb-object pairs in the job task descriptions and restricts his attention to patents identified as being related to robots, AI, or software. He uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings. By contrast, we infer document similarity by using geometric representations of word meanings (GloVe) that have been estimated directly from word co-occurence counts. Furthermore, we use not only the abstract but the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention. Rather than focus on specific technologies, we compute occupation-patent distance measures for all occupations and the entire set of USPTO patents since 1840.

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.1 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 appear likely to either change the way that an occupation performs its core work functions, or substitute for work done by that occupation.

We next consider some concrete examples of labor-displacive technologies. US patent number 6,289,319, titled "Automatic business and financial transaction processing system" 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.

Labor-displacive technologies can benefit new workers at the expense of incumbents. We next consider some specific examples of labor-displacive technologies discussed in Jerome (1934). We begin with 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 identified as breakthrough patents by Kelly et al. (2021). Both of these technologies benefitted skilled workers at the expense of unskilled labor.[7] In terms of related occupations, our methodology identifies various types of textile workers as being the some of the most relevant.

However, not all technologies benefit skilled labor. Goldin and Katz (2008) discuss how technological progress in manufacturing methods over the last two centuries involves two very distinct types of transitions. In an initial "deskilling" phase, production transitions from highly skilled artisans, each of which is involved with many stages of the production process, towards a "factory" model with more standardized tasks and extensive division of (mostly unskilled) labor (Atack et al., 2019).[8] In a second phase, manufacturing methods can become much more capital-intensive (e.g., via transitions towards continuous or batch processes) and explicitly labor-saving, so many, typically unskilled workers are replaced by fewer, typically more skilled technicians. Whereas the latter types of transitions – which Goldin and Katz (2008) note are likely more prevalent during the 20th century – are perhaps more displacive for low-skill workers, the opposite is likely to be the case for the former ones. A similar process likely plays out at the micro (e.g., task) level: *both* types of transitions likely occur regularly in the process of technological development, with both skilled and unskilled workers potentially being exposed to displacement risk, albeit probably for different reasons.

In brief, newer production methods may utilize a very different set of skills and expertise than older methods and current incumbent workers may lack the requisite skills. This displacement can take two forms: first, process innovation can lead to standardization of (formerly high-skill) tasks that now low skill can perform; second, the innovation may still require skill, but these are skills are different than the ones the incumbent workers possess. As concrete examples, 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); both patents are in the list of breakthrough patents identified by Kelly et al. (2021). 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, and wages of blowers and gatherers were reduced 40 per cent. In terms of our methodology, we identify "glaziers" and "molders, shapers, and casters, except metal and

[7] 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."

[8] Atack et al. (2019) write: "as useful as it is as an overall framing device, the Acemoglu and Restrepo model omits a fundamental feature of historical industrialization—namely, its extensive division of labor. As far as that model is concerned, the individual workers who perform tasks before and after automation could be the same people. In point of fact, *however, they were not the same people.* In the tiniest shops that are iconic depictions of hand production in early manufacturing, the artisan was highly skilled in the sense of performing most or all of the production tasks from start to finish, as well as 'nonproduction' tasks associated with managing the business. In the transition to machine labor, the artisan shop was displaced by the factory, which was different in many ways that could perhaps be summarized as 'more' of everything—more capital, more labor, and more output."

plastic" as being among the most related occupations to these two patents.[9]

Last, we also examine a more recent example of labor-displacive technology: the rise of e-commerce—and more specifically the automatic fulfillment of retail purchase orders. Our measure indicates the 1997 to 2002 period as 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.2 contains a longer list. Compared to all other clerk occupations, wage trends for order fulfillment clerks are fairly flat prior to 1997 (see Appendix Figure A.1). However, between 1997 and 2010, the wages of order clerks had declined by about 0.2 log points relative to all other clerks.

2.2 Relation to 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 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 $\mathcal{B}_t$ 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 sample.

Technology exposure and employment growth

We first relate our technology exposure measure (3) to employment growth. We use the Decennial Census surveys, which consist of repeated cross-sectional observations and contain information on occupations that we can link to our technology exposure measure. The data consists of an unbalanced panel of occupation–industry–decade employment shares and spans the Census decades

[9] Specifically, the latter occupation, which corresponds SOC code 519195, has a sub-occupation called "glass blowers, molders, benders, and finishers". Indeed as Jerome (1934) notes, glass workers displaced by the sheet and cylinder machines in their time were considered to be highly skilled workers: "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."

from 1910 to 2010. Appendix A.4 provides additional details.

We estimate the following specification,

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

Here, $Y_{i,j,t}$ is occupation i in industry j's share in total non-farm employment, so the dependent variable is the annualized percentage growth rate in employment. Given the secular changes in female labor force participation over the century-long time period, we restrict attention to male employment in this specification. Observations are weighted by the employment share of the given occupation–industry cell 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. We absorb industry-specific trends by including industry–time fixed effects in all specifications; depending on the specification, we include controls for the lagged 10-year employment growth rate.

Table 1 shows a strong and statistically significant negative correlation between our innovation measure η and subsequent changes in employment. The magnitudes are significant: a one-standard deviation increase in $\eta_{i,t}$ is associated with a 0.53 to 0.56 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 0.93 to 1.05 percentage point annualized decline—which corresponds to a cumulative decline in occupational employment of approximately 20 percent relative to unaffected occupations.

Technology exposure and wage growth

One concern is that the employment effects we document are amplified by the endogeneity of technological change: innovative effort is directed towards occupations for which labor supply is shrinking. Then the employment responses we are estimating include not only the effect of technology but also the fact that employment in these occupations would have contracted due to (future) labor scarcity. In this case, we would expect that our technology measure would predict wage increases in these occupations that are expected to experience scarcity of workers.

In contrast to this view, we find that our technology exposure measure predicts wage declines. To do so, we rely on data from the Current Population Survey Merged Outgoing Rotation Groups (CPS-MORG) which provides data on both wages and employment outcomes for the post-1980 period—see Appendix A.4 for details. We estimate a specification similar to (4), over horizons h 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 2 plots the estimated coefficients β along with 90% confidence intervals.

We see that our technology exposure 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. Employment growth also declines in this sample—approximately a 1.1% annualized decline in occupation employment over the next five to twenty years—which is quantitatively similar to the estimates reported in Table 1.

2.3 Is our technology exposure measure identifying mostly labor-displacive technologies?

The results in the previous sections suggest that our technology exposure measure is primarily identifying technologies that displace existing workers. Ex-ante, this fact is not obvious: the similarity between the description of an innovation and occupation tasks could in principle also capture technologies that enhance the productivity of incumbent workers. 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-displacive and innovations that enhance the productivity of incumbent workers.

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 (Cong, Liang, and Zhang, 2019). In brief, we extract the 500 most important topics from the text of patent documents and compute time-varying exposures of occupations to these patent topics. For each topic we estimate a version of equation (4) in the CPS-MORG data at the 10-year horizon, replacing our innovation exposure index $\eta_{i,t}$ with the topic k predictor. Last, we form an in-sample 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. Appendix A.6 contains further details.

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 $\eta_{i,t}$. That said, this finding is not particularly surprising given that the correlation between our baseline measure $\eta_{i,t}$ and these machine-learning predictors is over 70 percent. We conclude that our technology exposure measure captures labor-displacive innovations to a similar extent as a pure machine-learning predictor even though it is not a priori designed to do so.

2.4 Comparison to existing measures of technology exposure

We next explore whether our technology exposure measure contains additional information relevant in predicting wage and employment declines relative to existing measures of technology exposure, specifically the routine-task intensity (RTI) measure from Acemoglu and Autor (2011) and the Webb (2020) measures of exposure to robotics or software.

Given that these existing measures are purely cross-sectional in nature, we estimate a cross-sectional regression in long-difference similar to Webb (2020),

$$\frac{100}{h}\Big(\log Y_{i,j,2012} - \log Y_{i,j,1980}\Big) = \alpha + \alpha_j + \beta \eta^{\text{Pctile}}_{i,1980} + \delta X^{\text{Pctile}}_i + \epsilon_{i,j}, \tag{5}$$

where i indexes occupations and j indexes industries. In estimating (5), we combine information on wages and employment in the 1980 Census and the 2012 ACS data from Deming (2017). The dependent variable denotes either the log change in employment or the change in log wages over the 1980 to 2012 time period, hence $h = 32$. Since the Webb (2020) 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. We use the 1980 start-of-period percentile rank of our exposure measure. We include industry fixed effects α_j to account for industry specific shocks that may be correlated with occupational outcomes. Depending on the specification, the variables X_i include the to the routine-task intensity measure from Acemoglu and Autor (2011) or the Webb (2020) measures of exposure to robotics or software. We weight observations by the employment share in 1980 and cluster standard errors by occupation.

Table A.3 reports our findings for employment (Panel A) and wages (Panel B). Examining the table, we see that the point estimates of β are negative and highly significant across all specifications. Comparing the univariate specifications in columns (2) to (5), coefficient magnitudes on η are about 30 to 40 percent higher than the coefficient on RTI, 40 to 60 percent higher than the coefficient on the robot exposure measure, and more than double the coefficient on the software exposure. The estimated coefficients on our measure are only slightly attenuated when we include the Acemoglu and Autor (2011) or Webb (2020) measures. Focusing on the last column, we note that our measure of occupational exposure to technology contains independent information relative to the RTI measure, while it essentially subsumes the information in the Webb (2020) measures.

2.5 Descriptive Patterns in Technology Exposure

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

Which Occupations Are Most Exposed?

We find significant differences in the average degree of technology exposure across occupations. Table A.4 lists the top and bottom 25 occupations by average exposure over the entire 1850-2002 sample period for $\eta_{i,t}$. We see 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.

Autor and Dorn (2013) argue that recent patterns of job polarization—the disappearance of middle-skill (i.e., wage) occupations—are driven by their increased exposure to technological innovation relative to low- and high-wage occupations. Our direct measure of technology exposure confirms this view. Figure 3 plots the occupation technology exposures against average wage percentile ranks for the post-1980 period calculated using the CPS-MORG. Given the short time dimension of the data, we focus on cross-sectional comparisons. We note that the most exposed occupations tend to be found in the middle of the income distribution, consistent with the prevailing view regarding job polarization in the United States (Autor and Dorn, 2013; Bárány and Siegel, 2018).

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. We follow Acemoglu and Autor (2011) and focus on four task categories: manual tasks (routine and physical); non-routine manual (interpersonal); routine cognitive; and non-routine cognitive.[10] We then compute the average value of η, weighted by employment, across occupations that score in the top quintile of each of these categories; we then scale across the eight groups each year so that the total sums to one. Panel A of Figure 4 shows how the composition of which workers are most exposed has shifted across each of these task categories.

We see that for much of the sample, including the major innovation waves of the 1870 to 1890 and 1910 to 1930, occupations performing non-interpersonal manual tasks are on average most exposed. By contrast, occupations emphasizing cognitive tasks were 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. Over the last few decades, these occupations are almost as exposed to innovation as occupations emphasizing manual tasks. Given the composition of breakthrough innovations, this pattern is driven by the ICT revolution that has led to the modern digitalization of the workplace. Occupations that relate to these type of innovations have a distinctly

[10] Because the routine manual and non-routine manual (physical) task scores are highly correlated and also comove 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.

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 Deming (2017), who documents an increased importance of social skills in the labor market.

We next separate occupations by their education requirements. Specifically, we compute the share of workers in that occupation who have either completed a 4-year college degree or have attained a high-school diploma or lower in a given year. As above, we calculate the (weighted) average value of η across occupations in top and the bottom quintile of share of workers with or without a college degree each year. Panel B of Figure 4 plots the composition across these two categories.

We see that occupations requiring a college degree are, on average, significantly less exposed to innovation than occupations requiring a lower level of education. However, 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. As above, this pattern is driven by compositional shifts in the types of technologies being introduced, not changes in the share of workers with a college degree, since these categories are formed on purely cross-sectional comparisons.

3 Technology Exposure and Worker Outcomes

Armed with a measure of workers' technology exposure, we next move to the main goal of the paper: understanding how exposure to major technologies shapes outcomes of individual workers. The availability of wage earnings administrative data, combined with demographic information such as age, education or past earnings, allows us to not only understand the experience of individual workers but also how these experiences vary across workers.

3.1 Data

We combine survey information for a random sample of individual workers tracked by the Current Population Survey (CPS) Annual Social and Economic Supplement (ASEC) with employer-employee matched Social Security Administration (SSA) administrative data from their associated Detailed Earnings Records. These data include information on income and employer identification numbers (EINs) from Form W-2. The CPS ASEC includes information on occupation as well as demographic information such as age and gender. We complement the data with information on industry and average wages from the Census Longitudinal Business Database (LBD) using the EINs. We limit the sample to individuals who are older than 25 and younger than 55 years old and to periods where the CPS interview date is within the past 3 years so that the occupation information is relatively recent. Appendix A.5 contains further details.

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 follow Autor, Dorn, Hanson, and Song (2014) and Guvenen, Ozkan, and Song (2014) in that our main outcome variable of interest is the growth in worker i's average wage earnings, adjusted for life-cycle effects,

$$w^i_{t,t+h} \equiv \log\left(\frac{\sum_{j=0}^{h} \text{W2 earnings}_{i,t+j}}{\sum_{j=0}^{k} D(\text{age}_{i,t+j})}\right). \tag{6}$$

Here, W2 $\text{earnings}_{i,t}$ refer to the sum of all W2 income a worker i receives in a given year t.

Our data has detailed information on both the industry of a particular worker as well as the industry of origination of each patent. This allows us to exploit additional sources of variation: we can compare workers not only across occupations in the same industry, but also workers in the same occupation across industries. To this end, we build our technology exposure measure by also restricting attention to patents issued to firms in the same industry (4-digit NAICS) as the worker. Letting j index patents as before; $\mathcal{B}_{k,t}$ denote the set of breakthrough 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

$$\xi_{i,t} = \log\left(1 + \sum_{j \in \mathcal{B}_{k(i,t),t}} \rho_{o(i),j}\right). \tag{7}$$

In brief, our technology exposure metric (7) bears a strong similarity to (3), with the following modifications. First, our measure $\xi_{i,t}$ now varies by occupation, industry, and year instead of just occupation and year. Second, given that ξ 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. We have approximately 2.8 million person-year observations spanning the period from 1981 to 2016. Approximately 54% of the sample is male and 34% of the observations correspond to workers with a four-year college degree. The median worker in the sample is 41 years old and earns approximately \$50k per year in terms of 2015 dollars. The distribution of earnings is rather skewed: the average is equal to \$66k while the 5th and 95th percentiles are equal to \$16k and \$152k, respectively. Importantly, much of these differences in pay exist within industry–occupation groups; on average, about 58 percent of the cross-sectional dispersion exists within industry–occupation cells. To a large extent the same is true for education: approximately 56 percent of the variation in college share is within industry–occupation cells. These sizable differences in worker earnings within an industry–occupation cell are consistent with the view expressed in Goldin and Katz (2008) regarding the importance of within-job skill differences across workers. The top panel of Appendix Figure A.2 shows that the dispersion in wage earnings within an occupation and industry have been growing steadily over time, in line with the overall increase in income inequality across all workers.

These within industry–occupation differences in wage earnings are partly driven by firm heterogeneity: the workers that are paid the most relative to their peers in the same occupation and industry tend to work in more productive firms. That is, the firms that employ the workers at the top 5 percent of the within industry–occupation earnings distribution are approximately 0.45 log points more productive on average than the firms that employ the workers at the bottom 25 percent. However, the bottom panel of Appendix Figure A.2 shows that these between firm differences account for only a small fraction (about 16 percent) of the within industry–occupation earnings dispersion. See Appendix A.7 for further details on how we perform these calculations.

The last three rows of Table 3 report the distribution of the growth in lifecycle-adjusted cumulative earnings growth (6) across horizons. At a horizon of $h = 5$ years, the median is equal to approximately zero while the mean is -0.095; given that this variable 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.

3.2 Technology exposure and worker earnings growth

We estimate the following specification,

$$w^i_{t+1,t+h} - w^i_{t-2,t} = \alpha + \beta\,\xi_{i,t} + \mathbf{c}\,\mathbf{Z}_{i,t} + \varepsilon_{i,t}. \tag{8}$$

The dependent variable 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 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. Thus, our coefficient β is identified by comparing future earnings for an exposed worker to other workers in the same industry in a different occupation, or to workers in the same occupation in other industries. This saturated specification therefore allows us to partial out sources of time-series variation that occur at either the industry or the occupation level. In addition, we include flexible non-parametric controls for worker age and past worker earnings as well as recent earnings growth rates.[11] Standard errors are clustered at the industry level.

Overall, we find that workers' technology exposure is negatively related to their subsequent earnings growth. Panel A of Table 4 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 is the most conservative specification that compares a worker in a given year to either other workers in the same industry that year but in different occupations, or workers in the same occupation that year but in 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 0.023 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 magnitude 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.

These negative effects are comparable across workers in both manufacturing and services. Panel B of Table 4 compares the estimated coefficient β across workers employed in manufacturing and services, broadly defined. We see that 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 significant recent increase in the average technology exposure of workers in the service sector relative to manufacturing we saw in Section 2.5, strongly suggests that the displacement of workers in response to labor-displacive technologies is not purely a blue-collar worker phenomenon; rather it is increasingly present in white-collar occupations as well.

3.3 Worker Heterogeneity: Job Type, Education and Age

The negative average correlation between technology exposure and worker earnings can potentially mask considerable heterogeneity in outcomes across workers. Therefore, we next allow the estimated slope coefficient to vary by observable worker characteristics.

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

Job Type

First, we 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. 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 4.

Education

Existing work has emphasized that much of technological change is skill biased, where skill is often defined as worker education (see, e.g. Goldin and Katz, 2008). To this end, we next compare whether the earnings growth of workers with and without a four-year college degree respond differentially to the same increase in their occupation-industry technology exposure. This comparison exploits both between- as well as within-cell variation in the college share: recall that approximately 56 percent of the overall cross-sectional variation in the share of college educated workers is driven by within industry–occupation cell variation.

Panel A of Table 6 shows that there are 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.

Worker Age

Existing work has emphasized the specificity of human capital to particular technology vintages (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). If that is the case, we may expect that older workers, who have had more time to accumulate task-specific expertise, experience larger wage losses than younger workers in response to the same increase in technology exposure.

Consistent with this view, the bottom panel of Table 6 shows that older workers (those in the 45 to 55 range) experience significantly greater declines in earnings growth relative to younger workers (25 to 35 range) by approximately a factor of 1.8. This steep and monotonic gradient across age groups is consistent with the view that older workers are both more likely to have accumulated skills in existing technology and also less likely to become familiar with new production methods.

3.4 Worker Heterogeneity: Income

We next examine how the response of worker earnings growth to our technology exposure measure ξ 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 skill-biased, that is, complementary to worker skill. As such, we might expect that the most highly paid workers respond less—or even experience wage increases—as the technology that is relevant to their tasks improves.

However, our findings suggest a more nuanced view. Column (1) of Table 7 shows how the slope coefficient β varies 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 significantly greater wage earnings declines (more than twice the magnitude) than the average worker in response to a unit standard deviation increase in technology exposure ξ. There is some evidence that the earnings of lower-paid workers also respond more than average, but this difference is not statistically significant. 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. These findings reinforce our discussion in Section 2.1 that not all technological transitions–even skill-biased ones–need to disproportionately displace low-skill workers.

We next examine the extent to which these earnings declines are driven by job separation or unemployment spells. We do so by restricting the sample to workers that remain with the same firm over the next five years, or to workers that experience no significant period of unemployment—defined as the absence years without W2 income. Comparing columns (1) and (2) of Table 8, we see that firm exit is the key driver of earnings losses at the bottom of the income distribution; by contrast, it only accounts for less than 40 percent of the earnings losses for the highest-paid workers. Further, columns (3) and (4) show that these earnings losses are not driven by one- or three-year unemployment spells, respectively. Conditioning on workers that did not experience unemployment spells comparable patterns in earnings losses following an increase in technology exposure as the full sample of workers in column (1).

What type of worker heterogeneity is responsible for these income patterns? A number of additional results suggest that an important driver is variation in worker skill. Appendix Table A.5

explores alternative approaches for constructing income rankings. First, we drop workers that moved firms in the last year; this ensures that workers in the lowest bin are not workers that had partial employment. Comparing our baseline results in column (1) to column (2) we find that this has little impact on our results. In column (3) we instead rank workers based on their average income over the last two years; doing so leads to the same conclusions. We next verify that our effects are not driven by between-firm differences in average pay. 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 rank workers based on their wage earnings relative to the their employers' average wages (that is, the firms' total wage bill from the LBD divided by total employment). Comparing columns (4) and (1), we see the same income pattern, suggesting that employer heterogeneity is not the main driver of the income pattern we document.

We next remove heterogeneity in worker earnings unrelated to worker skill within an occupation–industry: we rank workers based on residual (log) wages net of occupation, industry, commuting zones, age, gender fixed effects—all interacted with calendar year. Removing commuting zone (interacted with year) effects allows us to remove the component of pay that may be related to local labor scarcity or monopsony power, while removing age bin by gender fixed effects accounts for the fact that older workers tend to be more highly paid. Column (5) in Appendix Table A.5 shows that ranking workers on these wage residuals has little impact on our income pattern of technology exposure. Last, in column (6) we also residualize wage earnings with respect to unionization status—interacted by calendar year. Doing so reduces the sample dramatically, since only approximately one in five workers in our sample provide an answer to the unionization question on the CPS. Even though the sample is smaller, we again see in column (6) the same pattern as before: the highest paid workers given these alternate definitions still experience greater wage earnings declines in response to an increase in technology exposure.

Further, we examine more broadly 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) of Appendix Table A.6, we see that the income gradient in Table 7 is present in industries with both high and low levels of unionization, though it is somewhat steeper 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 again dramatically smaller, we again find no meaningful differences in the response of the highest-paid workers to ξ as a function of union membership.

Last, we examine the degree to which these differences in earnings responses are driven by an 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 (8), 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 10th percentile of the earnings growth distribution across all workers in the same year. Column (1) of Table A.7 shows that top earners face significantly greater labor income risk than the average worker in response to an increase in their technology exposure. We see that 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, which is approximately three times higher than the average worker. In terms of magnitudes, this increase in tail risk accounts for less than half of the negative mean effect for the top earners, whereas it accounts for approximately two-thirds of the decline in earnings for workers at the bottom of the distribution. Column (2) shows that these effects are essentially driven by job separation. Conditioning the sample to workers that remain employed with the firm over the next five years, we see this increase in risk is essentially absent for workers that remain with their current employer.

In sum, we see that more highly-paid workers experience steeper relative earnings declines than the average worker following the introduction of new technologies that are related to the tasks they perform. Job separations—which are responsible for increases in the left tail of the earnings distribution—account for most of the earnings declines of the lowest-paid workers but only a moderate fraction of the losses of the highest-paid workers. Importantly, the underlying driver of this differential exposure as a function of past income is not the components of wages that are related to differences across industries, occupations, firms, geography (commuting zones), age, gender, or worker unionization status.

To the extent that this residual component of the worker's wage is related to her skill in performing her job, the fact that the more skilled workers experience larger losses following the introduction of new technologies suggests that some element of skill displacement is at work. This interpretation implies that the patterns in Table 7 are consistent with the standard view of technology as a complement to worker skill as long as 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 vintage.

Consistent with this view, we find that the income pattern we document is stronger in occupations in which human capital is both important and specific. In particular, we use data on occupation–related work experience requirements from O*NET to proxy for the amount of specific skills needed for a particular job (see Appendix A.5 for details). We then re-estimate equation (8) while also allowing the slope coefficient β to vary across occupations that score above or below the (worker-level) median in terms of required experience. Columns (1) and (2) of Table 9 show that in occupations in which related experience is important, 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, the difference between the highest- and lowest-paid group in the occupations in which related experience is less important is significantly smaller.

Last, Table 10 examines how the differential response of worker earnings across the earnings distribution varies with the characteristics of their job. There are two points worth noting here. First, the negative response of wage earnings at the bottom of the income distribution is significantly larger for occupations emphasizing routine cognitive and manual tasks. This suggests that the driving force displacing these workers is automation of certain tasks by technology. Second, and more importantly, the negative response of the most highly-paid workers (relative to their peers) is much more comparable across most task types—with the exception of high non-routine manual and interpersonal tasks. This suggests that the negative effect of the top is unlikely to be all driven by task automation; instead, these estimates are more consistent with these workers being less productive using the new technology, perhaps due to the lack of the necessary skills to effectively use the new technology—likely a more ubiquitous source of displacement risk for skilled workers.

3.5 Is all technology displacing workers?

Our findings so far might seem to suggest that there is a consistently negative effect of technology exposure on employment and wage growth. This conclusion is unwarranted: recall that we are only focusing on breakthrough technologies that are closely related to the tasks performed by workers. It is possible that other breakthrough technologies in the same industry as the worker that are unrelated to her tasks may complement her productivity.

In what follows, we focus on the breakthrough patents issued to firms in the worker's industry, that are *least likely* to be related to the tasks performed by a given worker. We construct this exposure measure in a similar fashion as (7),

$$\zeta_{i,t} = \log\left(1 + \sum_{j \in \mathcal{B}_{k(i,t),t}} \delta_{o(i),j}\right). \tag{9}$$

The key difference between (9) and (7) is that now δ is a measure of *distance* rather than similarity between the description of patent j and the tasks performed by occupation o.[12]

In sum, ζ now captures patents whose textual descriptions are unrelated to an occupation's task description. Given our findings thus far, we expect this measure to now capture patents that are

[12] We construct δ 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 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.

unlikely to displace the skills of a given worker. We therefore estimate a variant of (8)

$$w^i_{t+1,t+5} - w^i_{t-2,t} = \alpha + \beta\,\xi_{i,t} + \gamma\,\zeta_{i,t} + \mathbf{c}\,\mathbf{Z}_{i,t} + \varepsilon_{i,t}. \tag{10}$$

Table 11 reports the estimated coefficients β and γ from Equation (10). For brevity we focus on our preferred specification, which is saturated with both industry × year and occupation × year fixed effects, though results are similar with less saturated specifications.

The top panel of Table 11 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). Interestingly, the earnings of the workers at the bottom of the distribution increase about as much as those of the highest-paid workers in response to the same shock in ζ. Third, including ζ in this specification has very little influence on the estimated coefficients β relative to the specification in (8): 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.

3.6 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 the occupations that have higher requirements in terms of related experience than the median occupation.

Our interpretation of these patterns is that our measure picks up a combination of labor-saving innovations but also technologies that could be complementary to labor but may require skills that incumbent workers lack. Workers at the bottom of the earnings distribution are displaced as tasks are automated; job separations account for most of these earnings losses. Workers at the top of the earnings distribution are displaced because part of their accumulated skills (human capital) is specific to a particular technology vintage. As technology improves, part of these skills become obsolete. That is, even if the technology itself complements the skills of some workers, existing incumbents lack the necessary skills to effectively utilize the new technologies. Thus, workers who have accumulated the most skill in the existing technology also have the most to lose when new technologies arrive. Some of these earnings losses are driven by separations, but much of them are driven by lower earnings growth relative to their peers.

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. Though this is an interpretation that we cannot fully rule out—given that we do not directly observe labor productivity for individual workers—a number of our empirical findings are inconsistent with this mechanism being the key driver of the income gradient.

First, though it is certainly possible that variation in workers' bargaining power accounts for part of the dispersion in wage earnings within industry–occupation cells, we note that this dispersion is not only sizeable (accounting for 58 percent of the overall dispersion) but it has been increasing over time (Appendix Figure A.2)—unlike the rate of unionization or the labor share which has been declining over the same period (Karabarbounis and Neiman, 2013). Second, we find little evidence that this gradient varies across worker groups with varying degrees of bargaining power (rates of unionization). Third, comparing workers on wage residuals net of a number of characteristics that may affect wage earnings on top of worker skill, such as commuting zone or unionization status, leads to similar conclusions. Fourth, workers at the top of the income distribution are (somewhat) more likely to be employed at more productive firms: firms employing the highest-paid workers in an industry–occupation cell have approximately 45% higher labor productivity (0.45 log point units) than firms employing the workers at the bottom of the income distribution. Workers in these firms are less likely to have higher bargaining power relative to workers in less productive firms: Seegmiller (2021) finds that more productive US firms have lower labor shares, and face lower supply elasticities.[13] Fifth, job separations account for only a modest fraction of the wage earnings declines of the highest-paid workers in response to our technology exposure measure, while accounting for the entirety of the earnings response for the lowest-paid workers. This pattern is inconsistent with the idea that this income gradient is driven by firms reducing labor costs by replacing overpaid workers with labor-saving technologies.

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 literal 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

[13] See also Gouin-Bonenfant (2020) for related evidence among Canadian firms.

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. However, skills need not be transferable across technology vintages; equivalently, skill in an existing technology need not imply the automatic acquisition of skills necessary to use new technologies. In the next section we develop a simple model that explores this idea.

4 Model

Here, we develop a model with skill-biased technology in which worker skills are specific to a technology vintage. Workers perform two tasks, one of which is a complement and another which is a substitute to technology. Skilled workers perform tasks that are more highly remunerated and less likely to be automated—these tasks 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 if new technologies arrive that relate to tasks in which she has expertise.

4.1 Setup

The model applies to a single industry. Output in the industry 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

$$Y_t = \left[\mu\Big(H_t\Big)^{\sigma} + (1-\mu)\left(\lambda\Big(\xi_t\Big)^{\rho} + (1-\lambda)\Big(L_t\Big)^{\rho}\right)^{\sigma/\rho}\right]^{1/\sigma} \tag{11}$$

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 (11) 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,

$$\sigma < \rho < 1. \tag{12}$$

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

$$d\xi_t = -g\,\xi_t\,dt + \kappa\,d\,N_t. \tag{13}$$

Technology improves according to the Poisson process N_t with arrival rate $\omega\,dt$. We have set up output in per capita terms—thus, the negative drift term in equation (13) reflects the fact that population grows at rate g. Given (13), the level of ξ is stationary with long-run mean $\kappa\,\omega/g$.

When we map our empirical analysis to the model, we will interpret our 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—that is, the model's equivalent to 'occupations'. In particular, 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

$$H_t = \int_0^1 \theta\,\psi_t(\theta)\,d\theta, \tag{14}$$

where $\psi_t(\theta)$ is the measure of workers of skill level θ at time t. Since we normalize the total supply of labor to one, we have that $L_t = 1 - H_t$.

Workers vary in their ability to acquire skills in the new technology—increase their skill level θ. A subset of workers of measure s_l cannot acquire any skills, so they produce only in the low-skill task; for them $\theta = 0$. 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

$$\frac{d\theta_{i,t}}{\theta_{i,t^-}} = dM_{i,t} - h\,dN_{i,t}, \tag{15}$$

which we further assume to be reflected at the boundaries of the interval $[\underline{\theta}, 1]$.[14] In the above specification, $M_{i,t}$ is a doubly stochastic jump process, modeled as the product of a Poisson process with arrival rate ϕ that is independent across workers, times a random variable $m_{i,t}$ governs whether a worker learns or forgets; skills grow by a factor m with probability $1-z$ and shrink by a proportion 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, workers are displaced by the arrival of new technologies (improvements in ξ); this effect is captured by the last term in equation (15). The process $N_{i,t}$ is a doubly stochastic jump

[14] 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 $dM_{i,t}$ 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 $d\,N_{i,t}$ or $dM_{i,t}$ were to lead $\theta_{i,t}$ to fall below $\underline{\theta}$, we set the process immediately after the jump to $\underline{\theta}$.

process, which is however driven by the aggregate Poisson process N_t—describing improvements in technology in (13). Specifically,

$$dN_{i,t} = d_{i,t}\, dN_t. \tag{16}$$

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. 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.

The aggregate supply of skilled labor H_t in (14) 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

$$w_{i,t} = W_{L,t} + \theta_{i,t} \left(W_{H,t} - W_{L,t} \right). \tag{17}$$

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

$$W_{H,t} = \frac{\partial Y_t}{\partial H_t}, \quad \text{and} \quad W_{L,t} = \frac{\partial Y_t}{\partial L_t}. \tag{18}$$

4.2 Model Calibration

We discuss the model calibration next.

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 13 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.

The next feature of the data to target is the response of labor productivity (11) and the labor share to improvements in technology ξ. This response is informative, since an increase in ξ in the model has an ambiguous impact on both output and the labor share. The sign and the magnitude of the response depend on the extent to which different tasks contribute to output (μ and λ); technology-labor complementarity (ρ and σ) and the response of H and L to a technology shock—recall that the aggregate supply of high- and low-skill inputs H and L varies in the short run, due to skill displacement (15).

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 (15), we also target the heterogeneity in earnings responses to changes in technology (see Section 3). Specifically, we target the mean earnings growth responses and changes in worker tail risk—column (1) in Table 7 and column (4) in Table A.7. 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,

$$\frac{w_{i,t+h} - w_{i,t}}{w_{i,t}} = \underbrace{\frac{w_{l,t}}{w_{l,t} + \theta_{i,t} s_{p,t}}}_{\substack{\text{low skill} \\ \text{income share}}} \underbrace{\frac{\Delta_h w_{l,t+h}}{w_{l,t}}}_{\substack{\text{low skill} \\ \text{wage chg}}} + \underbrace{\frac{\theta_{i,t} s_t}{w_{l,t} + \theta_{i,t} s_t}}_{\substack{\text{high skill} \\ \text{income share}}} \left[\frac{s_{p,t+h}}{s_{p,t}} \cdot \underbrace{\frac{\Delta_h \theta_{i,t+h}}{\theta_{i,t}}}_{\substack{\text{skill} \\ \text{displacement}}} + \underbrace{\frac{\Delta_h s_{t+h}}{s_t}}_{\substack{\text{high skill} \\ \text{wage chg}}} \right] \tag{19}$$

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 (19), whether the highest-earning workers experience larger declines depends on whether the increase in the skill premium is sufficient to offset the loss of worker skill θ due to skill displacement—see equation (15). 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 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 (that is, the technology is related to her occupation), while separately controlling for exposure and shock dummies, analogous to our empirical specification that includes industry-time FE. To be consistent with our empirical work, we include income-bin specific intercepts and exposure coefficients. 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 the parameter vector Θ by minimizing the distance between the output of the model $\hat{X}(\Theta)$ and the data X,

$$\hat{\Theta} = \arg\min_{\Theta} \left(X - \hat{X}(\Theta) \right)' W \left(X - \hat{X}(\Theta) \right). \tag{20}$$

We choose the weighting matrix W to obtain a balanced fit across the target moments. See

Appendix A.8 for details on the model solution and calibration.

Table 12 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.

Examining Table 13, 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.

Discussion of the Mechanism

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, while the supply of low-skill labor L increases. This fall in H is temporary, the supply of high-skill labor 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 even as output increases. 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.

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. $d_{i,t} = 0$). 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 responses is the second column in Table 11. 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. Part of the reason for this divergence is that labor supply is exogenous in the model—whereas adjusting hours worked may be important for the lowest-paid workers.

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 generates 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 considerably 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 (plotted in orange) 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.

Taking stock, our model illustrates that our main empirical results are consistent with the prevailing view that recent technological advances are complementary to the human capital of skilled workers, once we take into account the possibility that some skills are specific to earlier vintages. This exercise highlights the distinction between the skill premium and income inequality in the model: income inequality depends not only on skill prices, but also the quantity of skill workers have, which can change even in the short run as new vintages of technology displace skills specific to prior vintages. Thus, changes in technology affect both quantities and prices, hence movements in the skill premium are insufficient to fully characterize earnings inequality.

5 Conclusion

We develop a methodology for identifying the arrival of labor-displacive technologies that relies only on the textual description of the patent document and the tasks performed by workers in an occupation. Upon further inspection, our measure likely picks up a combination of labor-saving innovations as well as technologies that could be complementary to labor but may require skills that incumbent workers lack. 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.

Individual workers experience significant losses in wage earnings following improvements in related technologies: the average worker experiences approximately a 0.02 log point decline in her cumulative wage earnings over the next five years. Importantly, we find significant heterogeneity in these responses across age and income levels. In particular, older and the more highly-paid workers are significantly more affected than the average worker. Our interpretation of these patterns is that workers at the bottom of the earnings distribution are displaced as tasks are automated; job separations account for most of these earnings losses. Workers at the top of the earnings distribution are displaced because part of their human capital is specific to a particular technology vintage (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). 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. Some of these earnings losses are driven by separations, but much of them are driven by lower earnings growth relative to their peers. We quantitatively replicate these findings in the context of a model that features skill-biased technical change and vintage-specific human capital.

References

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

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

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

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

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

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

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

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

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

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

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

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.

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.

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

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

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

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

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

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

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

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

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

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

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

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*.

Ruggles, S., C. A. Fitch, R. Goeken, J. D. Hacker, M. A. Nelson, E. Roberts, M. Schouweiler, and M. Sobek (2022). IPUMS ancestry full count data: Version 3.0.

Ruggles, S., S. Flood, R. Goeken, M. Schouweiler, and M. Sobek (2022). IPUMS USA: Version 12.0.

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

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

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

Figures and Tables

Figure 1: Occupations Most Related to Three Breakthrough Innovations

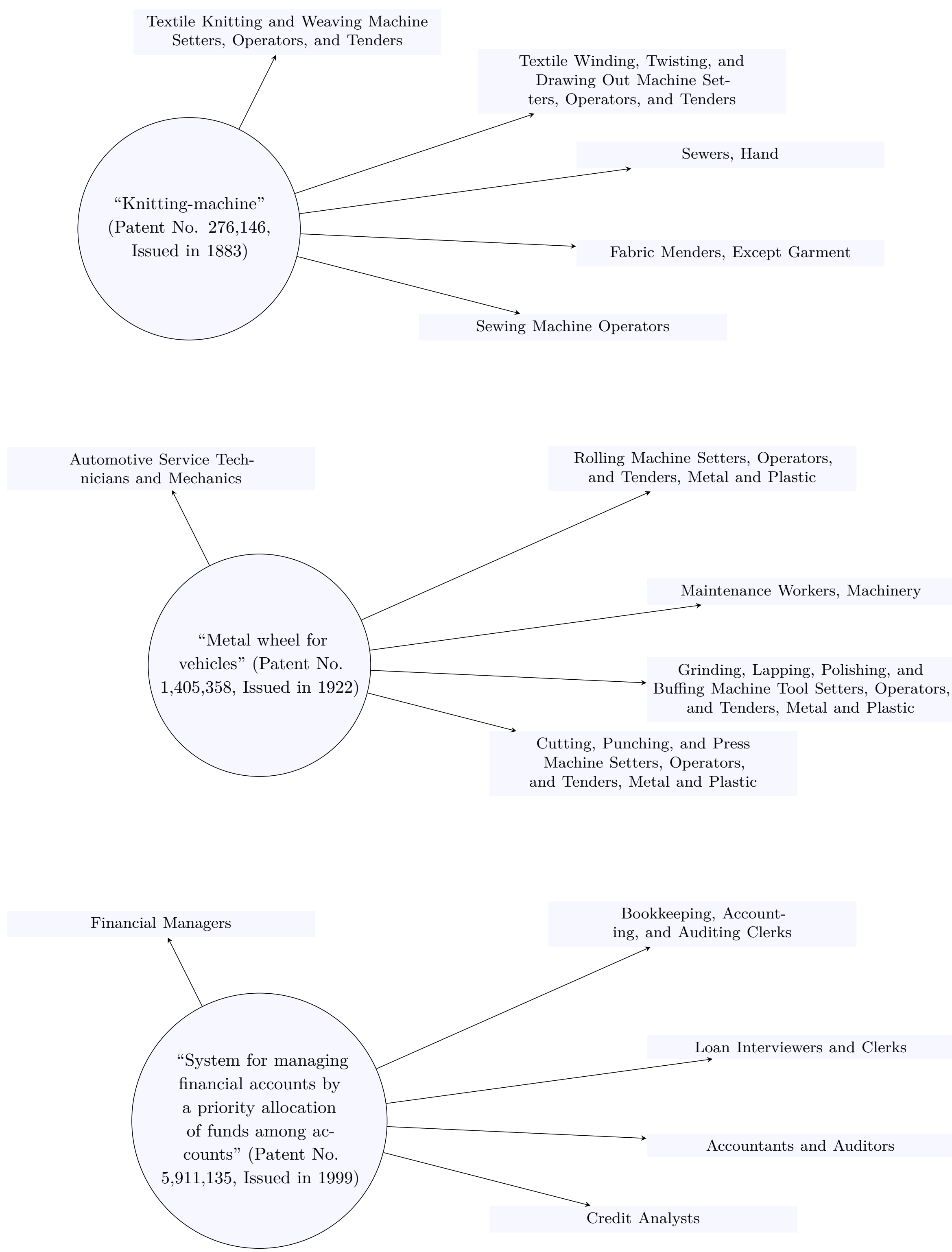
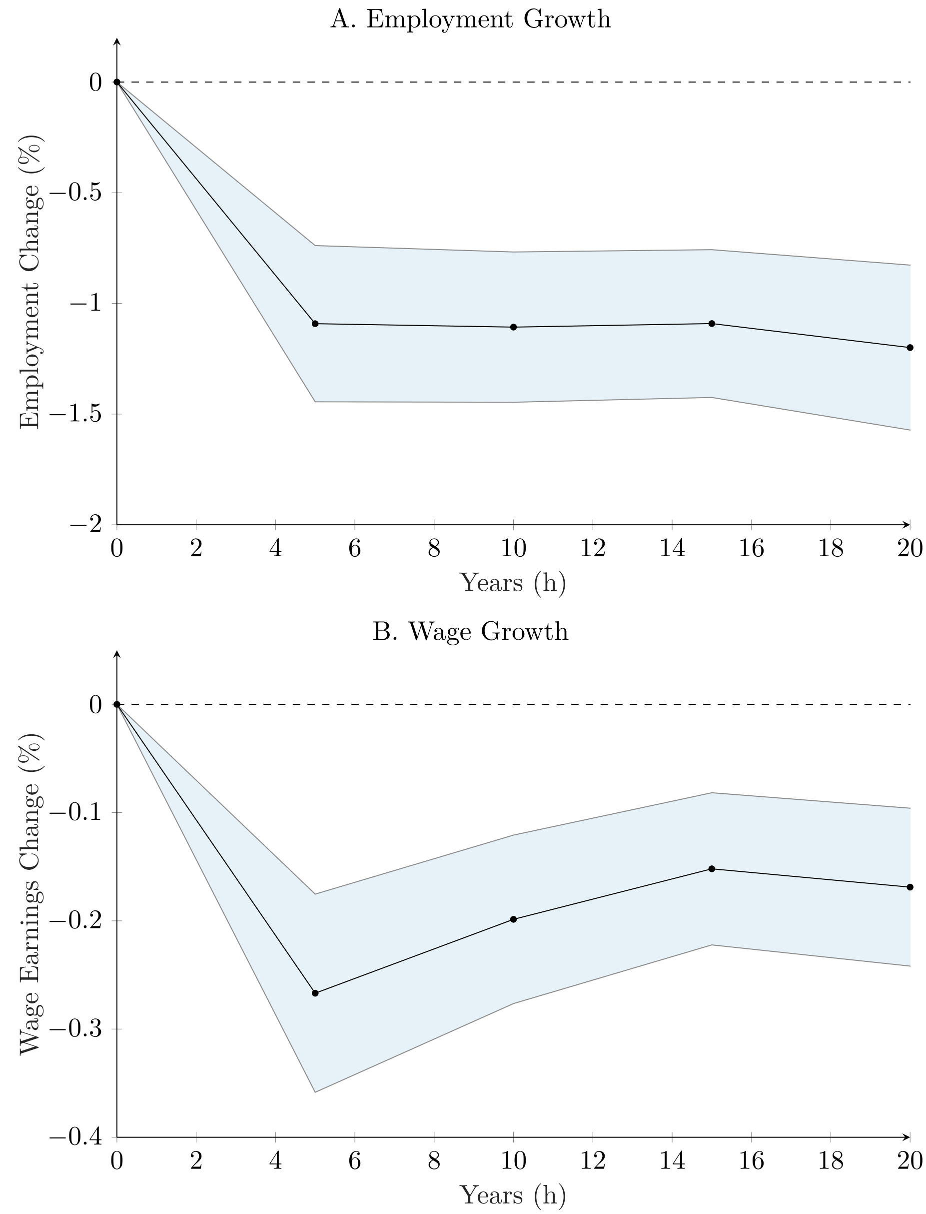


Figure 2: Employment, wage earnings and technology exposure (recent period: 1980–present)

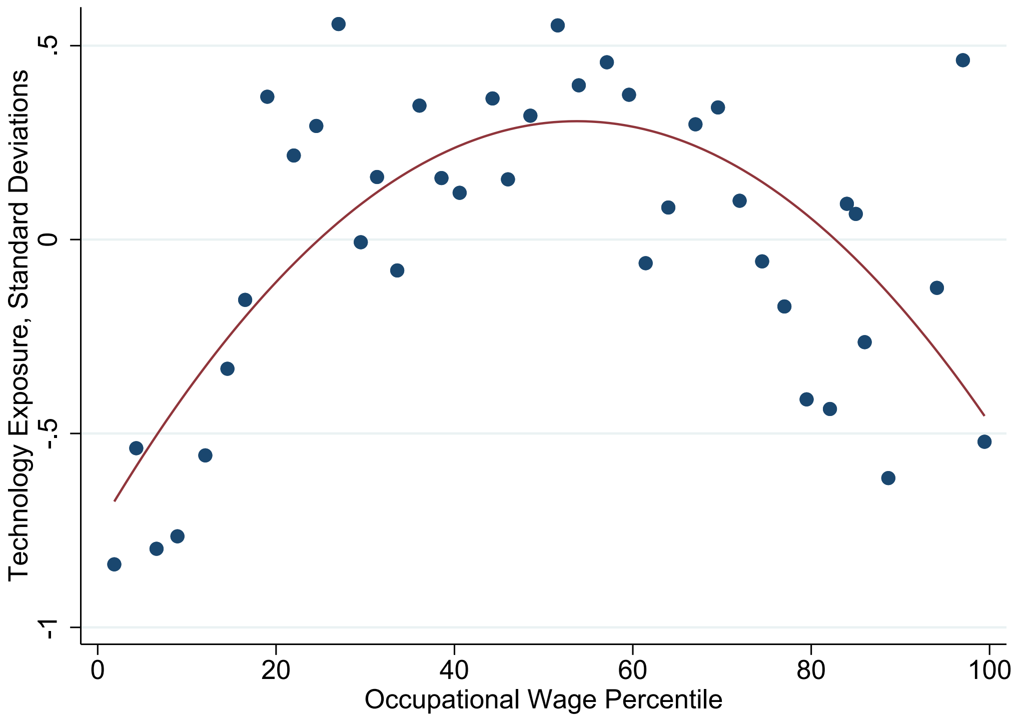


Note: Figure plots coefficients from panel regressions of annualized wage and income growth rates over different time horizons on occupation innovation exposures:

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

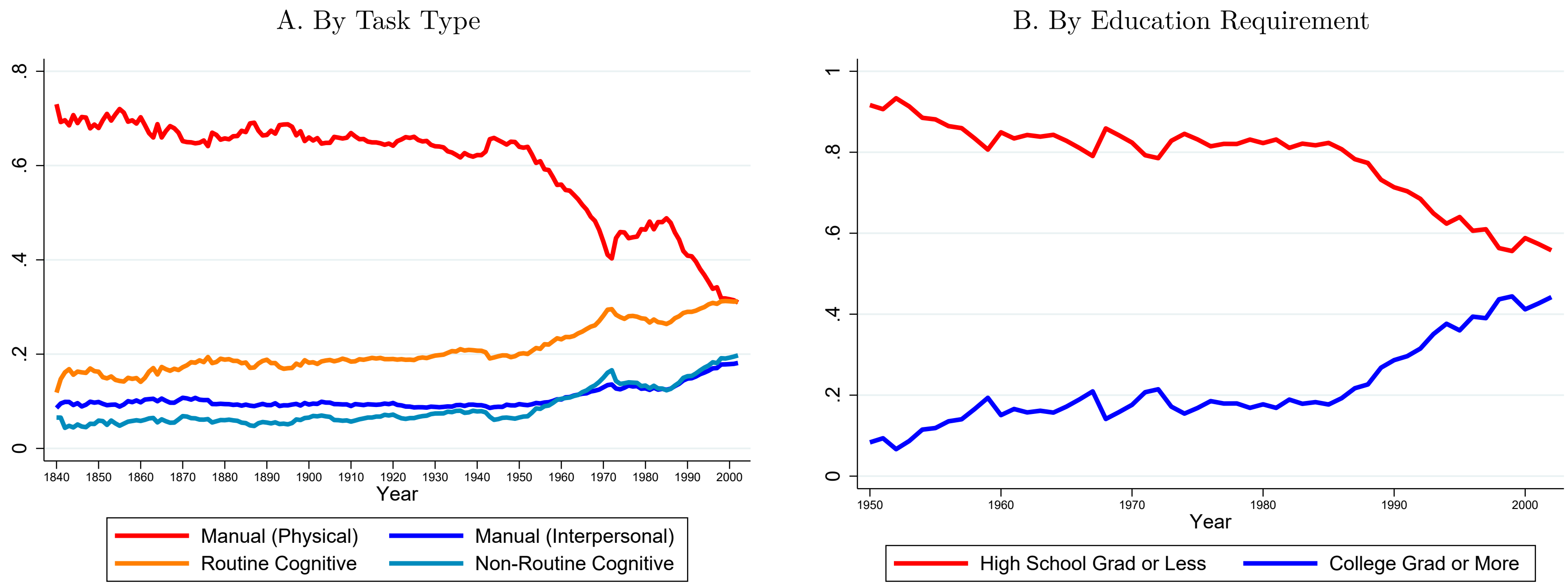
Here Y_i denotes occupation employment share or wage. Controls $X_{i,t}$ include three one-year lags of dependent variable, and time fixed effects. Dependent variable is expressed in annualized percentage terms and $\eta_{i,t}$ is standardized. Figures plot 90% confidence interval for each time horizon. Data come from the CPS Merged Outgoing Rotation Groups (MORG) and cover the 1983–2018 period.

Figure 3: Technological Exposure, by occupation income level



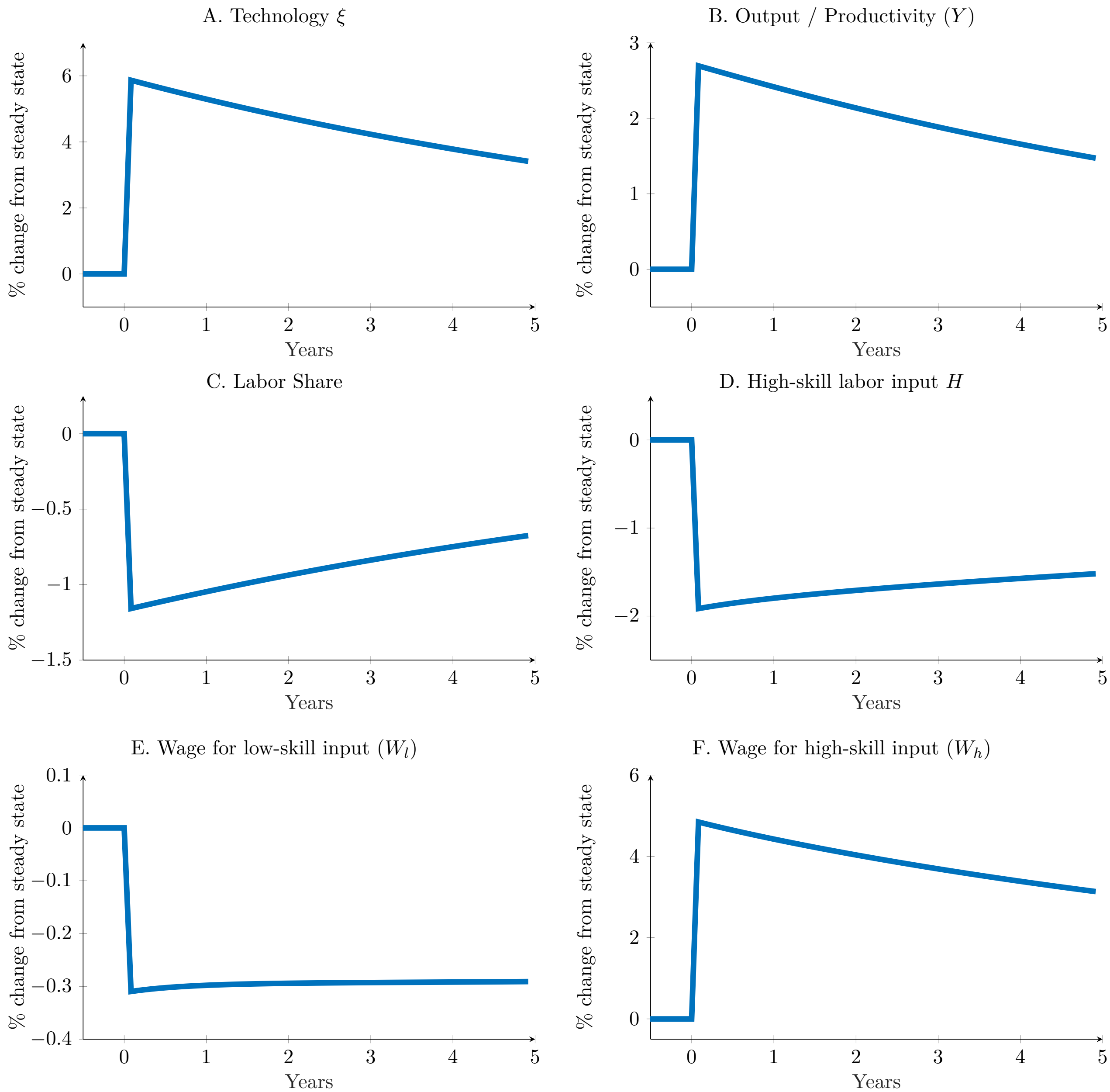
Note: This figure plots a bin scatter for technology exposure $\eta_{i,t}$ (in standard deviation units) by occupational wage percentile. The period covers the years 1979 to 2002, and occupations are sorted according to wage percentile rank within each year. The wage data come from the Current Population Survey Merged Outgoing Rotation Groups.

Figure 4: Technology exposure, composition across occupation categories



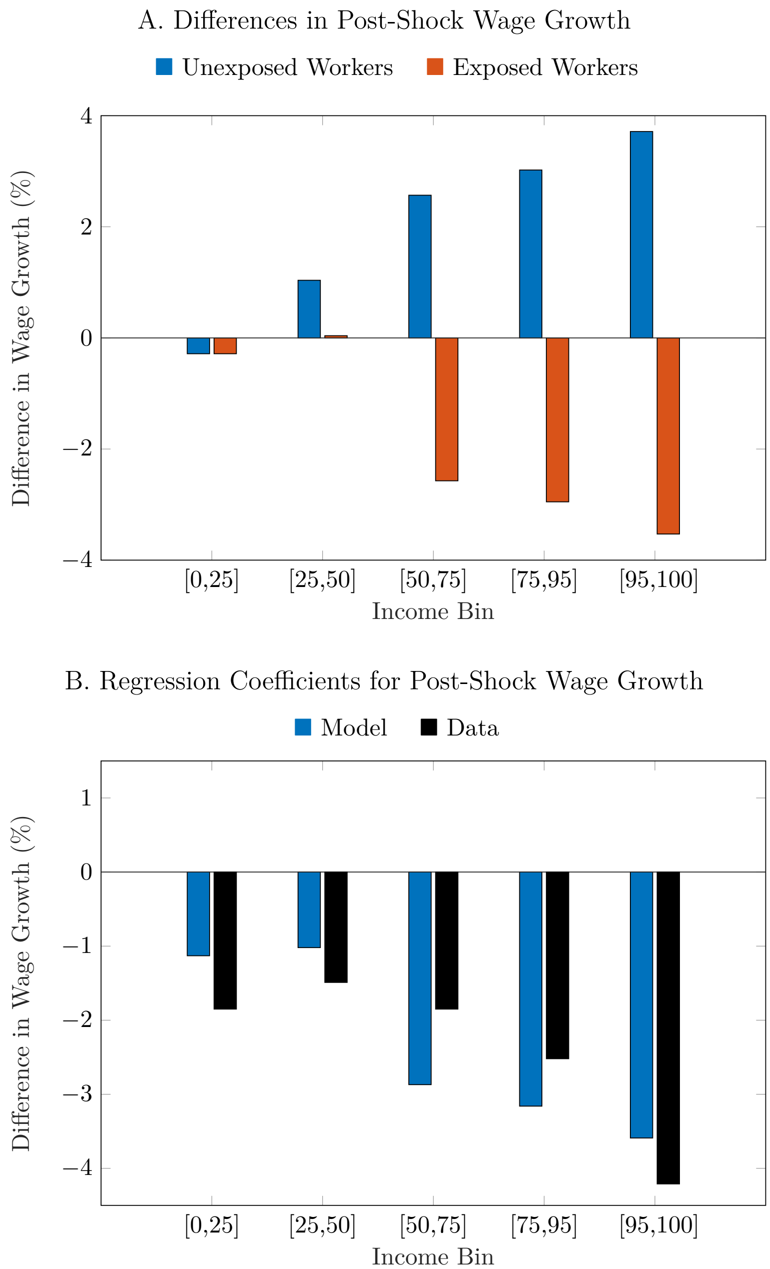
Note: Panel A of the figure plots the composition of our index of technological exposure by task category w. Specifically, we calculate, $\lambda_{w,t} = \sum_i \eta_{i,t} \times T_{w,t}(i) \times \omega_i$, where $T_{w,t}$ is an indicator that takes a value of 1 if occupation i is in the top quintile 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". $S_{s,t}(i)$ is an indicator that takes a value of 1 if occupation i is in the top quintile of the cross-section distribution of shares of workers in category s at time t. 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.

Figure 5: Model: Impulse Responses



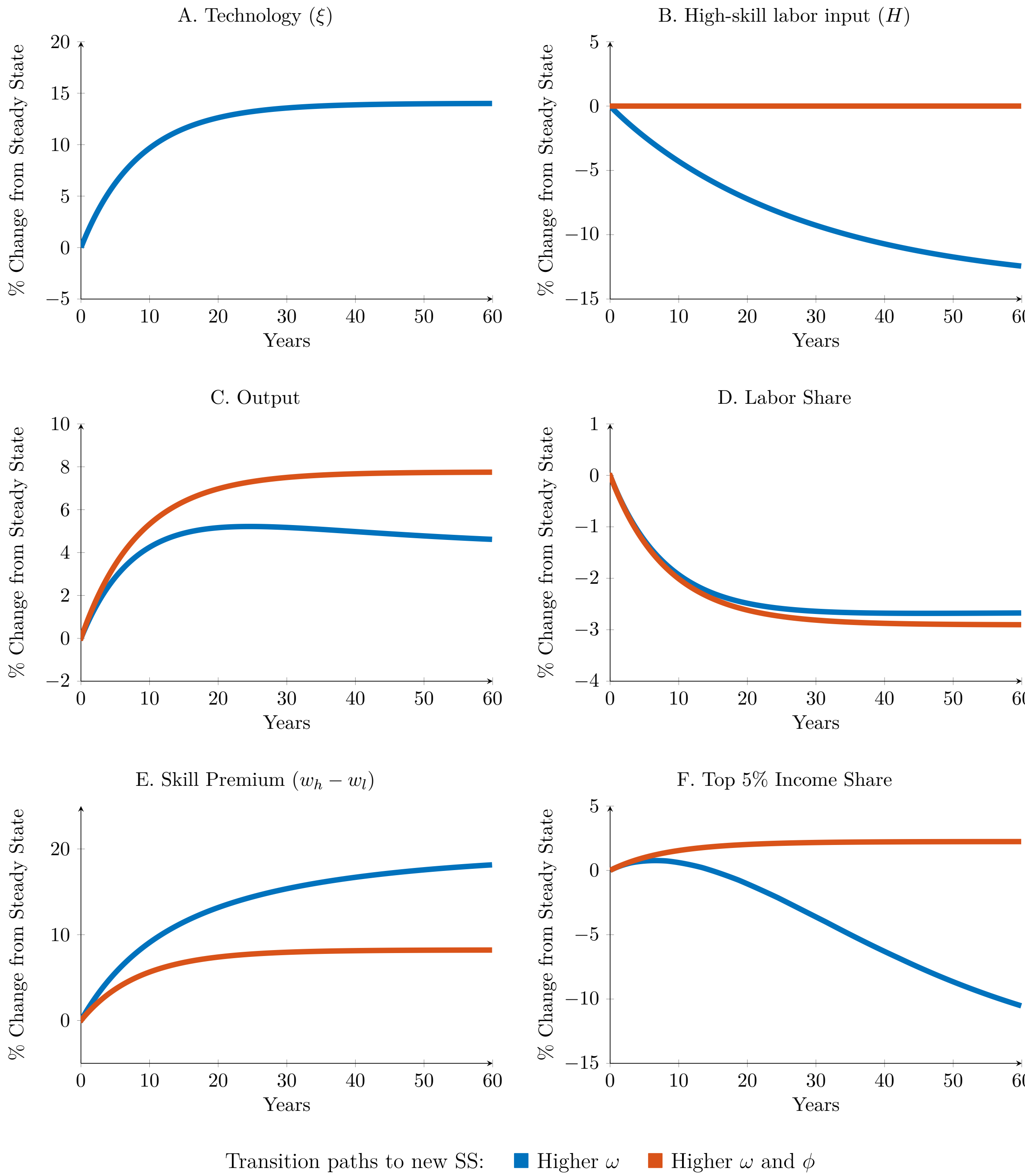
Note: This figure shows the impulse responses of key model quantities following a one-standard deviation technology shock evaluated at the steady state of the model.

Figure 6: Model: Innovation and Worker Earnings



Note: Panel A shows raw differences between wage growth during a shock period and wage growth during a non-shock period for workers who are exposed to the shock versus 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 transition paths from 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 plots the resulting transition path for ξ); 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 plots the corresponding transition paths of the quantity of high-skill labor input; Panel C plots total output/productivity; panel D plots the labor share of output; panel E plots the skill premium; and panel F plots income inequality, defined as the top 5% income share in the model.

Table 1: Technology And Employment Over the Long Run (1910–present)

Employment Growth	Horizon			
	10 Years	20 Years	10 Years	20 Years
Technology Exposure, $\eta_{i,t}$	-0.53	-0.93	-0.56	-1.05
(past decade average)	(0.17)	(0.23)	(0.18)	(0.23)
Observations	102,400	81,009	72,451	54,662
Controls				
Industry × Time FE	Y	Y	Y	Y
Lagged Dependent Variable			Y	Y

Note: We report estimated slope coefficients β from the following regression

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

for $h = 10, 20$ years for Census years spanning from 1910-2010. Here $Y_{i,j,t}$ is the occupation-industry cell i, j share in total male non-farm employment. Tech exposure $\eta_{i,t}$ is standardized and growth rates are in annualized percentage terms. We report standard errors clustered at the occupation level in parentheses beneath coefficient estimates. Observations are weighted by 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	-1.116	-0.191
	(0.206)	(0.048)
B. Machine learning predictors of employment declines		
Mean across topics that predict employment declines	-1.180	-0.187
	(0.172)	(0.032)
First PC across topics that predict employment declines	-1.059	-0.185
	(0.164)	(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. We report standard errors clustered at the occupation level in parentheses beneath coefficient estimates. 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 1983–2018 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	40.0	7.6	29	30	33	40	46	51	53	632,000
Age, workers in top-5 income bin	43.2	6.8	31	33	38	44	49	52	53	109,000
Occupation-industry technology exposure (ξ)	0.647	0.971	0	0	0	0.232	0.866	2.028	2.922	1,495,000
Lifecycle-adjusted earnings growth, 3-years	-0.072	0.478	-0.973	-0.507	-0.129	0.008	0.129	0.312	0.472	2,773,000
Lifecycle-adjusted earnings growth, 5-years	-0.095	0.526	-1.116	-0.622	-0.174	0.002	0.142	0.337	0.507	2,596,000
Lifecycle-adjusted earnings growth, 10-years	-0.145	0.609	-1.363	-0.825	-0.280	-0.023	0.160	0.388	0.576	1,697,000
Male	0.542	0.498	0	0	0	1	1	1	1	2,782,000
Has four-year college degree	0.344	0.475	0	0	0	0	1	1	1	2,782,000
Exit firm within 1 year	0.175	0.380	0	0	0	0	0	1	1	2,676,000
Exit firm within 5 years	0.516	0.500	0	0	0	0	1	1	1	2,210,000
Union member	0.165	0.271	0	0	0	0	0	1	1	553,000
Industry unionization rate	0.163	0.163	0	0	0.031	0.111	0.250	0.436	0.475	2,398,000
Log revenues per worker	5.048	1.106	3.227	3.691	4.468	5.106	5.741	6.375	6.745	966,000

Note: The table reports summary statistics for our wage earnings data from the Census Detailed Earnings Record (DER)-CPS merged sample, which covers the 1981 to 2016 period. The sample includes all workers whose unique identifiers (PIK codes) can be matched between the DER and CPS data for CPS years between 1981 and 2016 and who satisfy labor force attachment sampling criteria. W2-Earnings are reported in terms of 2015 dollars. The occupation-industry technology exposure ξ is defined as in (7) from the main text. Patents are matched to industry of origination using information from the confidential Census SSL and LBD datasets. The variable "Has four-year college degree" denotes whether a given individual has completed a 4-year degree at the time they were observed in the CPS. 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 (6) in the main text. Log revenues per worker come from the Longitudinal Business Database revenues file and have coverage from 1997 to the end of our sample. Earnings growth rates, technology exposure, and log revenues per worker are all winsorized at the 1% level each year. 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 × Year		x		x
Industry × Year			x	x

Note: Table shows the estimated slope coefficients β (times 100) from equation (8) in the main text. The dependent variable is workers' cumulative earnings growth (net of life-cycle effects) over the next 3, 5, or 10 years. We report standard errors clustered at the industry (NAICS 4-digit) level in parentheses 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. Prior income rank bins are based on workers' yearly earnings rank within their occupation-industry pair. To define occupation boundaries we continue to use David Dorn's revised Census occ1990 occupation codes. For industries we use the 4-digit NAICS code of a worker's primary employer, with the exception that when there are fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification when assigning income ranks. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th and 95th percentile; and, 95th percentile and above.

Table 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 (8) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary by occupation task type intensity. 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 in parentheses 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—in addition to task type dummies. 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; each task type row in the table corresponds to a separate regression. The right column reports the p-value in percent units from a two-sided test that the coefficients are equal for above- and below-median task intensity interactions.

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 × Year		x		x
Industry × Year			x	x

Note: Table shows the estimated slope coefficients β (times 100) from equation (8) 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 in parentheses 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, 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 (8) 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 in parentheses 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 are fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification when assigning income ranks. Within these groups we partition workers into the following earnings bins: between bottom and 25th percentiles; between 25th and median; between median and 75th percentile; between 75th and 95th percentile; and, 95th percentile and above. The bottom panel of the table reports differences within each column between the top and bottom income bin coefficients, the top income bin and average of the middle three income bins, and bottom income bin and average of the middle three income bins, along with corresponding p-values (in percent units) for a two-sided test of coefficient equality.

Table 8: Technology exposure and worker earnings growth, conditioning on continuing employment

Worker earnings rank (rel. to occ × ind)	All Workers	Stay in Firm	Continual Employment	
		Same EIN for at least next 5 yrs	ExitSR=0 no years with missing W2s	ExitLR=0 no 3-cons years with missing W2s
0–25th percentile	-1.85	-0.49	-1.20	-1.49
	(0.49)	(0.41)	(0.37)	(0.45)
25–50th percentile	-1.49	-1.03	-1.22	-1.36
	(0.43)	(0.31)	(0.33)	(0.41)
50–75th percentile	-1.85	-1.35	-1.35	-1.57
	(0.39)	(0.27)	(0.29)	(0.35)
75–95th percentile	-2.52	-1.79	-1.96	-2.26
	(0.43)	(0.29)	(0.32)	(0.37)
95–Top	-4.21	-2.62	-3.55	-3.94
	(0.59)	(0.43)	(0.47)	(0.51)
95–Top vs 25–95th percentile	-2.26	-1.22	-2.04	-2.21
(p-val, %)	0.00	0.25	0.00	0.00
95–Top vs 0–25th percentile	-2.36	-2.12	-2.35	-2.45
(p-val, %)	0.10	0.07	0.00	0.00
0–25th percentile vs 25–95th percentile	0.10	0.90	0.31	0.24
(p-val, %)	77.70	0.57	22.73	42.54

Note: Table shows the estimated slope coefficients β (times 100) from equation (8) 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 in parentheses 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. Column (2) focuses on workers who remain with the same firm (EIN) over the next five years. Column (3) removes workers reporting no W2 earnings in any year in the next 5 years, and column (4) omits from the sample 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. The bottom panel of the table reports differences within each column between the top and bottom income bin coefficients; top income bin and average of the middle three income bins; and, bottom income bin and average of the middle three income bins, along with the corresponding p-values (in percent units) for a two-sided test of coefficient equality.

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

Dependent Variable:	Earnings Growth	
Related Experience Requirement:		
	Low	High
0–25th percentile	-2.88	-1.16
	(0.55)	(0.50)
25–50th percentile	-1.96	-1.14
	(0.50)	(0.45)
50–75th percentile	-2.16	-1.61
	(0.43)	(0.42)
75–95th percentile	-2.22	-2.65
	(0.49)	(0.45)
95–Top	-3.71	-4.38
	(0.70)	(0.74)
95–Top vs 25–95th percentile	-1.60	-2.58
(p-val, %)	2.61	0.00
95–Top vs 0–25th percentile	-0.83	-3.22
(p-val, %)	35.35	0.01
0–25th percentile vs 25–95th percentile	-0.77	0.64
(p-val, %)	5.09	10.22

Note: Table shows the estimated slope coefficients β (times 100) from equation (8) in the main text, where 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 in parentheses 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 are fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification when assigning income ranks. The bottom panel of the table reports within-related experience category differences between the top and bottom income bin coefficients; top income bin and average of the middle three income bins; and, bottom income bin and average of the middle three income bins, along with corresponding p-values (in percent units) for a two-sided test of coefficient equality.

Table 10: Technology exposure and worker earnings growth by income rank, comparison across occupations emphasizing different task types

Worker earnings rank (rel. to occ × ind group)	Manual Physical		Routine Cognitive		Non-Routine Cognitive		Non-routine Manual/Personal	
	Low	High	Low	High	Low	High	Low	High
0–25th percentile	-1.10	-2.48	-1.21	-2.27	-2.33	-1.43	-1.86	-1.58
	(0.56)	(0.54)	(0.58)	(0.52)	(0.52)	(0.52)	(0.50)	(0.60)
25–50th percentile	-0.71	-2.06	-0.53	-2.11	-1.59	-1.35	-1.71	-0.70
	(0.49)	(0.51)	(0.46)	(0.48)	(0.47)	(0.46)	(0.45)	(0.57)
50-75th percentile	-0.92	-2.56	-0.67	-2.52	-2.11	-1.58	-2.40	-0.22
	(0.47)	(0.44)	(0.43)	(0.45)	(0.42)	(0.43)	(0.43)	(0.51)
75-95th percentile	-1.87	-2.84	-1.49	-2.93	-2.68	-2.36	-3.08	-0.55
	(0.54)	(0.45)	(0.51)	(0.42)	(0.45)	(0.49)	(0.43)	(0.58)
95–Top	-3.72	-4.36	-3.69	-4.18	-4.84	-3.71	-4.95	-2.29
	(0.81)	(0.73)	(0.68)	(0.63)	(0.57)	(0.86)	(0.59)	(0.78)

Note: Table shows the estimated slope coefficients β (times 100) from equation (8) in the main text, where coefficients on technology exposure $\xi_{i,t}$ are allowed to vary with occupation–industry earnings rank and occupation types that score above or below the median in terms of the Acemoglu and Autor (2011) task types. We report standard errors clustered at the industry (NAICS 4-digit) level in parentheses beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. See notes to Tables 5 and 7 for further details.

Table 11: 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 (10) 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 in parentheses 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 are fewer than 10 workers in such an occupation-industry-year we move to the broader 2-digit NAICS industry classification when assigning income ranks. The bottom panel of the table reports differences within each column between the top and bottom income bin coefficients; top income bin and average of the middle three income bins; and, bottom income bin and average of the middle three income bins, along with corresponding p-values (in percent units) for a two-sided test of coefficient equality.

Table 12: 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: Table reports the parameters 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 13.

Table 13: Model Fit

Statistic	Data	Model
Labor share, response to ξ	-1.25	-1.16
Skill premium (p75 / p25 ratio), average	2.01	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 13.

A Appendix

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

Here, we briefly overview our conversion of unstructured patent text data into a numerical format suitable for statistical analysis. We obtain text data for measuring patent/job task similarity from two sources. Job task descriptions come from the revised 4th edition of the Dictionary of Occupation Titles (DOT) database. We use the patent text data parsed from the USPTO patent search website in Kelly et al. (2021), which includes all US patents beginning in 1976, comprising patent numbers 3,930,271 through 9,113,586, as well as patent text data obtained from Google patents for pre-1976 patents. Our analysis of the patent text combines the claims, abstract, and description section into one patent-level corpus for each patent. Since the DOT has a very wide range of occupations (with over 13,000 specific occupation descriptions) we first crosswalk the DOT occupations to the considerably coarser and yet still detailed set of 6-digit occupations in the 2010 edition of O*NET.[15] We then combine all tasks for a given occupation at the 2010 O*NET 6-digit level into one occupation-level corpus. The process for cleaning and preparing the text files for numerical representation follows the steps outlined below.

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

Stop words come from the following sources:

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

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

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

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

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.

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

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

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 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.[16] 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 overweighing rare occurrences or extremely frequent occurrences. 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

[16] 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.).

that the basis for this word vector space is arbitrary and has no meaning; distances between word embeddings are only well-defined in relation to one another and a different training instance of the same data would yield different word vectors but very similar pairwise distances between word vectors.

The next step consists of using these word vectors to construct measures of document similarity. To begin, we first construct a weighted average of the word embeddings with a document (a patent or occupation description). Specifically, we represent each document as a (dense) vector X_i, constructed as a weighted 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. We also note that there is not a Pennington et al. (2014) word embedding estimate for every word in every patent. For example, occasionally there are OCR text recognition errors in patent documents, which can lead to misspelled words that do not have an embedding estimate as a result; there may also be patents that use extremely specific and technical terms (i.e. chemical patents introducing an new and very specific molecule) which do not have a word embedding available. In such cases we simply calculate (A.6) for the set of words which do have a Pennington et al. (2014) embedding estimate—which is the vast majority of terms in practice.

In natural language processing, a common approach to emphasize terms that are most diagnostic of a document's topical content is the 'term-frequency-inverse-document-frequency' (TF-IDF). In brief, $TFIDF_{i,k}$ overweighs word vectors for terms that occur relatively frequently within a given document and underweighs terms that occur commonly across all documents. We follow the same approach: in constructing (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. 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 (2020), who also analyzes the similarity between a patent and O*NET job tasks. Webb (2020) focuses on similarity in verb-object pairs in the title and the abstract of patents with verb-object pairs in the job task descriptions and restricts his attention to patents identified as being related to robots, AI, or software. He uses word hierarchies obtained from WordNet to determine similarity in verb-object pairings. By contrast, we infer document similarity by using geometric representations of word meanings (GloVe) that have been estimated directly from word co-occurence counts. Furthermore, we use not only the abstract but the entirety of the patent document—which includes the abstract, claims, and the detailed description of the patented invention. In addition to employing a different methodology, we also have a broader focus: we are interested in constructing time-series indices of technology exposures. As such, we compute occupation-patent distance measures for all occupations and the entire set of USPTO patents since 1836.

Last, as we describe in the paper, we perform several adjustments to the raw measure of similarity (2). First, we remove yearly fixed effects. We do so in order to account for language and structural differences in patent documents over time.[17] 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 $\rho_{i,j}$ for all patent–occupation pairs, we restrict to the set of patents identified as breakthroughs by the KPST procedure when we construct our time-varying measures η and ξ of occupational and occupation-industry exposure to technology—as in the main text equations (3) and (7), respectively. Additionally, though we initially calculate $\rho_{i,j}$ for DOT documents with occupations defined at the 2010 SOC code level, we compute η and ξ using the modified Census occ1990 scheme (commonly referred to as "occ1990dd" codes) from Autor and Dorn (2013). The occ1990dd codes are comparable over time and can be readily linked to the CPS or Census, which proves useful for our analysis of occupational- and worker-level outcomes across various datasets. After crosswalking SOC occupations to the occ1990dd level, we take the average of the original 2010 SOC version of $\rho_{i,j}$ within each linked occ1990dd occupation code; we then use the resulting occ1990dd-level measure of $\rho_{i,j}$ in equations (3) and (7).

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

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

To determine the set of breakthrough patents that are relevant to a given industry, KPST map patents to industries based on their Cooperative Patent Classification (CPC) technology class using the probabilistic mapping constructed by Goldschlag et al. (2020). Here, $\alpha_{I,p}$ denotes the probability of breakthrough patent p being assigned to industry I. We compute $\alpha_{I,p}$ using the Goldschlag et al. (2020) 3-digit CPC class to 6-digit NAICS probabilities, aggregated up to the 4-digit NAICS level. 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

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

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. We include in the vector of controls $\mathbf{Z}$ year fixed effects and the lagged 5-year growth rate of the outcome variable.

[17] 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.

Appendix Figure A.3 plots the estimated impulse response coefficients $\beta(h)$ along with corresponding 90% confidence intervals. Panel A illustrates that an one-standard-deviation increase in the degree of innovation $\psi_{I,t}$ 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 $\psi_{I,t}$, 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, consistent with aggregate labor-technology substitution (Acemoglu and Restrepo, 2018).

A.4 Census public-use data

Decade-level Census Data

We gather Census data from IPUMS and compute aggregate employment shares for occupations in Census years spanning 1910-2010 (data for 2010 come from the American Community Survey). For the 1910, 1920, 1930, and 1940 Census years we obtain complete count Census datasets (Ruggles, Fitch, Goeken, Hacker, Nelson, Roberts, Schouweiler, and Sobek, 2022), while we use the default IPUMS (Ruggles, Flood, Goeken, Schouweiler, and Sobek, 2022) samples for each decade from 1950 to 2010. We make use of occupation definitions that IPUMS indicates are the most time-series consistent. For pre-1950 Census years we employ the 1950 Census occupation coding scheme, since the more modern revised 1990 Census classification scheme is only available for Census years 1950 and later. We use the revised 1990 Census occupation classifications in the years they are available in IPUMS. 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. Because not all occ1990dd codes crosswalk to a 1950 Census code, this results in an unbalanced panel at the occ1990dd $\times$ Census decade level.

When we compare employment outcomes across Census decades, we keep the same underlying Census occupation coding scheme as the basis for crosswalking to the occ1990dd level. For example, when looking at 20-year employment growth from 1940 to 1960, we crosswalk occupations in 1940 and 1960 to the occ1990dd level by way of the 1950 Census occupation codes; however, for 1950-1970 employment growth we use the revised 1990 Census occupation codes for 1950 and 1970, since the 1990 version of occupation codes come available in 1950. Finally, when we run regressions analyzing employment growth within an occupation-industry cell, we define industries using the 1950 Census industry designations, which are available the furthest back in time. Because the distinction between Census industry and occupations are less reliable before 1910 we start our analysis using the data from the 1910 Census.[18] We measure employment by occupation and industry within a Census decade by summing up IPUMS sample weights ("perwt") for males between the ages of 20 and 60 who report being in the labor force in that Census year.[19]

[18] The first Census year to offer separate entries for activities now designated as industry and occupation is 1910. See `https://usa.ipums.org/usa/volii/91occtc.shtml`.

[19] We use labor force participation because employment status is not observable in the 1920 Census and the definition of employment status is less comparable over the first few decades in our sample (see `https://usa.ipums.org/usa-action/variables/EMPSTAT#comparability_section`).

Annual Current Population Survey (CPS) Data

We use the Current Population Survey Merged Outgoing Rotation Groups (MORG) to analyze annual outcomes for repeated cross sections at the occupation level in the post-1980 period. We obtain the cleaned versions of MORG extracts provided by the Center for Economic Policy Research (CEPR).[20] Using these data we construct a time series of wage and employment growth for occupations at the occ1990dd level starting in 1979 with a crosswalk from occ1990dd to yearly CPS codes. For hourly wages we use “rw", the CEPR-preferred variable for analyzing hourly wages in the post-1979 period, which combines the usual hourly earnings for hourly workers and non-hourly workers and adjusts for top-coding using a lognormal imputation; we take the average of hourly wages within an occupation, weighted by the CPS earnings weight (“orgwgt"). To measure employment we sum up earnings weights by occupation for individuals who report being employed in a given CPS year. Because the CPS extracts use the Census 1970 coding scheme prior to 1983, and only a subset of occ1990dd codes match to a Census 1970 occupation code, we start our time-series analysis of employment and wage changes in the post-1983 time period (i.e. Table 2 and Figure 2) when these extracts began using the 1980 Census occupation classification scheme. The last CPS year in our sample is 2018.

A.5 Census administrative data

DER-CPS Sample

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

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

Longitudinal Business Database

We merge the DER-CPS sample with the Longitudinal Business Database (hereafter LBD, Jarmin and Miranda (2002)) to get worker industry information and data on employers' wages and revenues. We assign workers to a consistently-defined industry code based on their employer's industry codes in the LBD. To do this we use a deduplicated crosswalk from LBD firm identifiers (“firmid") to firm EINs, and we assign each EIN with the Fort and Klimek (2018) 2012 4-digit NAICS code based off LBD industry information.

[20] Available here: https://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-data/

We compute firm-level wages by aggregating yearly LBD payroll and employment information to the EIN level, and we compute firm average wages for a given year by taking the ratio of payroll to total employment. Since LBD annual payroll information is reported for the last 12 months as of March of that year, while W2 earnings are reported over the calendar year, we use the average of the LBD-implied wage for years t and year $t+1$ as a measure of employers' average wages in year t. We use the 2016 of edition of the LBD to compute firm average wages and to obtain firm industry information. Finally, we utilize the newer revenue-enhanced LBD (Haltiwanger, Jarmin, Kulick, Miranda, and Penciakova, 2019) to measure revenue productivity, defined as revenues per worker; for similar reasons as for firm wages, we take the average of revenues per worker in years t and $t+1$ to proxy for revenue productivity in year t. We have LBD information on firm industries and wages from the start of our sample in 1981 through 2015, while revenue information is available from 1997 through 2016.

Assigning Industry of Origination to Patents

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

Restrictions for Assigning Worker Earnings Bins

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

Related Experience Measure

Some of our analysis of worker-level outcomes in the DER-CPS sample makes use of an occupational related experience measure; we outline the construction of this metric here. 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, and merge the resulting measure with our DER-CPS worker-year panel. We then sort workers into high- and low-related experience occupations based on whether they are above or below the individual worker-level cross-sectional median in terms of occupational years of related experience required.

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 $\mathcal{B}_t$. 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

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

As before we only sum over breakthrough patents and we 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 occupations 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 Groups sample (1983-2018) used in the employment regressions in Figure 2. We estimate the set of topics for patents issued in the post-1976 period. We restrict attention to this period for two reasons: first, this overlaps with 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, giving us an upper bound on the extent of such mixing.

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 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.

A.7 Details on Worker Earnings Variance Decompositions and Analysis of Labor Revenue Productivity

In this appendix section we provide details on calculations discussed in section 3.1 of the main text; these include a variance decomposition of worker earnings into between- and within-occupation-industry/firm type components and correlating firm revenue productivity with worker earnings rank.

Earnings Decompositions

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

$$\overline{\log\left(\text{W2 Earnings}_{i,t}\right)} = \alpha_{occ(i,t)\times ind(i,t),t} + \epsilon_{i,t} \tag{A.16}$$

Here i indexes an individual worker and $\alpha_{occ(i,t)\times ind(i,t),t}$ denotes yearly fixed effects for worker i's occupation-industry pair at time t, and $\overline{\log\left(\text{W2 Earnings}_{i,t}\right)}$ denotes cross-sectionally demeaned log earnings. Similar to the restrictions we impose for assigning earnings rank bins, we use a worker's occ1990dd $\times$ 4-digit NAICS code for $\alpha_{occ(i,t)\times ind(i,t),t}$ if there are 10 or more observations in the yearly occ-ind cell, and 2-digit NAICS code if fewer than 10 observations. We drop an occ-ind cell if there are still fewer than 10 observations after moving to 2-digit NAICS codes. Estimating (A.16) results in the following variance decomposition of annual log earnings:

$$\text{Var}_t\left(\overline{\log\left(\text{W2 Earnings}_{i,t}\right)}\right) = \underbrace{\text{Var}_t\left(\hat{\alpha}_{occ(i,t)\times ind(i,t),t}\right)}_{\text{Between occupation-industry}} + \underbrace{\text{Var}_t\left(\hat{\epsilon}_{i,t}\right)}_{\text{Within occupation-industry}} \tag{A.17}$$

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

Next we examine an extension on the decomposition in (A.16) by adding dummies for firm type:

$$\overline{\log\left(\text{W2 Earnings}_{i,t}\right)} = \alpha_{occ(i,t)\times ind(i,t),t} + \alpha_{\text{FirmRank}(i,t),t} + \epsilon_{i,t} \tag{A.18}$$

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

$$\begin{aligned}\text{Var}_t\left(\overline{\log\left(\text{W2 Earnings}_{i,t}\right)}\right) = &\underbrace{\text{Var}_t\left(\hat{\alpha}_{occ(i,t)\times ind(i,t),t}\right)}_{\text{Between occupation-industry}} + \underbrace{\text{Var}_t\left(\hat{\alpha}_{\text{FirmRank}(i,t),t}\right)}_{\text{Between firm-type}} + \\ &\underbrace{2\times\text{Cov}_t\left(\hat{\alpha}_{occ(i,t)\times ind(i,t),t}, \hat{\alpha}_{\text{FirmRank}(i,t),t}\right)}_{\text{Sorting by occ-ind and firm-type}} + \underbrace{\text{Var}_t\left(\hat{\epsilon}_{i,t}\right)}_{\text{Within occupation-industry and firm-type}}\end{aligned} \tag{A.19}$$

In panel B of A.2 we graph the time series of the earnings decomposition in (A.19). The sample for this panel ends in 2014 instead of 2015 since we average the LBD-implied firm average wage over two adjacent years (as explained in appendix A.5), and our LBD firm wage coverage cuts off in 2015. The residual component $\text{Var}_t\left(\hat{\epsilon}_{i,t}\right)$ now represents the variance of earnings within both occupation-industry and firm type; the variance of this within component again increases throughout the sample period. The full period average share of variance of the within occupation-industry and firm type component is now 48.4%; this implies that accounting for firm type effects and sorting can only explain about 16.1% (i.e. 1 - 48.4/57.7) of the within occupation-industry variance share of log earnings.

We also report a similar variance decomposition for college attainment by estimating a version of (A.16), where a dummy for being a college grad replaces worker earnings as the dependent variable. We find that 56.1% of variation in college attainment can be explained within occupation-industry cells over the full sample period.

Revenue Productivity of Employers and Worker Earnings Ranks

Next we describe in detail our calculations from section 3.1 on how worker earnings rank within occupation-industry correlates with firm labor productivity. In particular, we estimate the following:

$$\log\left(\text{Revenue/Worker}_{i,t}\right) = \alpha_{EarningsRank(i,t)} + \alpha_{occ(i,t)\times ind(i,t),t} + \epsilon_{i,t} \tag{A.20}$$

We use the revenue-enhanced version of the LBD to calculate $\log\left(\text{Revenue/Worker}_{i,t}\right)$, the log revenues per employee of individual i's year-t employer. The term $\alpha_{EarningsRank(i,t)}$ denotes dummies for worker i earnings rank at time t. These correspond to the same earnings bins (i.e. [0%, 25%), [25%, 50%), [50%, 75%), [75%, 95%), [95%, 100%]) as throughout the paper. We also control for yearly occupation-industry fixed effects so that comparisons are within occupation and industry. We use the [0%, 25%) bin as the omitted category. While our main sample spans 1981-2016, the sample period for this regression is constrained by availability of LBD revenues data begins in 1997. The estimates among the top four income bins are as follows: 0.122 (SE = 0.012); 0.231 (SE=0.019); 0.345 (SE=0.026); and, 0.454 (SE=0.034) for the [25%, 50%), [50%, 75%), [75%, 95%), and [95%, 100%] income rank bins, respectively.[21] Hence we find that workers in the highest occupation-industry income bin tend to be employed at firms with about 0.45 higher log revenue productivity than workers in the bottom bin, and this estimate is highly statistically significant. Looking at the distribution of revenue productivity in Table 3, this corresponds to a difference of about 0.4 standard deviations. The revenue productivity effects are also increasing in worker earnings rank across all income bins, suggesting a strong monotonic relationship across the earnings distribution.

A.8 Model Appendix

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

Setup

For convenience, we repeat some of the assumptions of the model from section 4 of the main text here. Production is given by a nested CES production function involving three inputs: technology ξ, high-skill labor H, and low-skill labor L. In the inner nest, a composite good X is produced via a combination of L and ξ

$$X_t = \left(\lambda \xi_t^{\rho} + (1-\lambda) L_t^{\rho}\right)^{(1/\rho)}.$$

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

$$Y_t = \left(\mu H_t^{\sigma} + (1-\mu) X_t^{\sigma}\right)^{(1/\sigma)}.$$

Technology evolves with process

$$d\xi_t = -g\xi_{t^-}\,dt + \kappa\, d\,N_t$$

where dN is a Poisson random variable with annualized intensity ω capturing the arrival of new breakthrough technologies and κ increase in technology associated with a new breakthrough. ξ_t is closely related to an Ohrnstein-Uhlenbeck process, except that it has compound Poisson rather than Brownian innovations.

Wages associated with H (w_h) and L (w_l) are calculated as the marginal product of each task. Given

[21] We report standard errors clustered by industry.

output, these are calculated as

$$W_{H,t} = \mu H_t^{\sigma-1}(\lambda\xi_t^{\rho} + L_t^{\rho}(1-\lambda))^{(\sigma/\rho)} + \mu H_t^{\sigma})^{(1/\sigma-1)},$$

and

$$W_{L,t} = (1-\mu)(1-\lambda)L_t^{\rho-1}(\lambda\xi_t^{\rho} + (1-\lambda)L^{\rho})^{(\sigma/\rho)} + \mu H_t^{\sigma})^{(1/\sigma-1)}(\lambda\xi_t^{\rho} + (1-\lambda)L_t^{\rho})^{(\sigma/\rho-1)}.$$

As is noted in equation 17, a worker's wage is a weighted average of w_H and w_L.

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

The model features a continuum of workers, each with a state parameter θ on the [0, 1] interval, corresponding to a worker's ability to produce the high-skill output H. Share s_h of workers have the ability to accumulate H over time, which follows a stochastic process we describe further below, while fraction $1-s_h$ of workers have $\theta = 0$ throughout their careers. Workers die with Poisson arrival rate δ, and are replaced by newborn workers. Among these workers, s_h begin with the lowest skill level $\underline{\theta}$ and can accumulate human capital, and all remaining workers have the absorbing $\theta = 0$ state. Without loss of generality, we index the unit interval so that those who can accumulate skills have $i \leq s_h$.

Conditional on survival and $i \leq s_h$, worker i's human capital evolves according to

$$d\theta_{i,t} = m_{i,t}\,\theta_{i,t^-}\,dM_{i,t} - d_{i,t}\,h\,\theta_{i,t^-}dN_t. \tag{A.22}$$

In this case, $dM_{i,t}$ is a random variable representing human capital acquisition, m is the size of the jump relative to initial human capital, dN_t 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 a worker is exposed to a technology shock ($d_{i,t} = 1$), she experiences a h% human capital loss, otherwise she does not.

Computing model moments analytically

To perform analytical approximations characterizing both aggregate quantities as well as model analogues to our cross-sectional regressions, we must characterize and integrate over the distribution of θ. To do this, we discretize the model with a monthly time-step $\Delta = 1/12$. To do this, we first discretize the support of θ over an exponentially increasing, finite grid of points on the [0, 1] interval. As noted in the text, we impose upper and lower reflecting barriers at $\underline{\theta}$ and 1, respectively. We choose the θ grid so that θ increments are m log points apart. So, learning/forgetting moves a worker exactly up or down by 1 grid point. When h isn't an integer multiple of m, we instead assume that the worker falls $\lfloor h/m \rfloor$ or $\lceil h/m \rceil$ gridpoints, where we choose the probabilities of the two different outcomes so that $\log\theta$ drops by h on average. Since the model features a continuum of workers, the law of large numbers holds exactly, and we can thus work directly with expectations when solving the model.

The aggregate state of the economy is $\mathbf{\Omega_t} \equiv [\xi_t, \pi_t']$, where π_t summarizes the fraction of workers with skill at each grid point $\in [\underline{\theta}, \ldots, 1)$.[22] Given the assumed stochastic structure and our discretization, the aggregate state Ω can be represented has (2 state) Markov-Switching VAR representation, ie a VAR with parameters that differ across the breakthrough ($\chi_t \equiv N_t - N_{t-\Delta} = 1$) and no-breakthrough states $\chi_t = 0$

$$\mathbf{\Omega_t} = c_{\chi_t} + A_{\chi_t}\mathbf{\Omega_{t-\Delta}}, \tag{A.23}$$

where A_{χ_t} is a matrix of state-dependent VAR coefficients and c_{χ_t} is a vector of state-dependent constants. This is easy to see if we consider first the discretized law of motion for ξ_t:

$$\xi_t = (1-\Delta g)\xi_{t-\Delta} + \chi_t \cdot \kappa,$$

so the first elements of c_0 and c_1 are 0 and κ, respectively, and the 1,1 elements of both A_0 and A_1 are both $1 - \Delta g$. While we abstract away from details associated with imposing reflecting barriers, notice that in the interior of the support of θ, the vector of probabilities evolves $\pi_{\mathbf{t}}$ according to the law of motion:

$$\begin{aligned}
\pi_{i,t} \quad = \quad & \underbrace{(1-\delta)}_{\substack{\text{adj for}\\ \text{worker exit}}} \Big[\pi_{i,t-\Delta} - \underbrace{\phi\Delta\, \pi_{i,t-\Delta}}_{\substack{\text{outflow from}\\ \text{learning/obsolescence}}}\Big] + \underbrace{\phi\Delta(1-z)(1-\delta)\,\pi_{i-1,t-\Delta}}_{\substack{\text{inflow from}\\ \text{learning}}} + \underbrace{\phi\,\Delta\, z\,(1-\delta)\pi_{i+1,t-\Delta}}_{\substack{\text{inflow from}\\ \text{skill obsolescence}}} \\
+ \quad & \chi_t(1-\delta)\alpha \underbrace{\left[\tfrac{h/m - \lfloor h/m \rfloor}{m}\pi_{i+\lfloor h/m \rfloor, t-\Delta} + \tfrac{\lceil h/m \rceil - h/m}{m}\pi_{i+\lceil h/m \rceil, t-\Delta} - \pi_{i,t-\Delta}\right]}_{\text{inflow/outflow from skill displacement}}.
\end{aligned}$$

Collecting coefficients together allows us to define a recursive system for probabilities which evolves according to:

$$\begin{bmatrix} \pi_{\mathbf{t}} \\ s_h \cdot (1 - \mathbf{1}'_{\mathbf{dim}(\pi)}\pi_{\mathbf{t}}) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{\Phi_{11}}_{,\chi_{\mathbf{t}}} & \mathbf{\Phi_{12}}_{,\chi_{\mathbf{t}}} \\ \mathbf{\Phi_{21}}_{,\chi_{\mathbf{t}}} & \mathbf{\Phi_{22}}_{,\chi_{\mathbf{t}}} \end{bmatrix}}_{\equiv \mathbf{\Phi}_{\chi_{\mathbf{t}}}} \begin{bmatrix} \pi_{\mathbf{t-\Delta}} \\ s_h \cdot (1 - \mathbf{1}'_{\mathbf{dim}(\pi)}\pi_{\mathbf{t-\Delta}}) \end{bmatrix}, \tag{A.24}$$

where $\mathbf{1}_{\mathbf{dim}(\pi)}$ is a column vector of ones with the same dimension as $\pi_{\mathbf{t}}$. $\mathbf{\Phi}_{\chi_{\mathbf{t}}}$ is the transpose of the discretized Markov transition matrix for an individual, potentially skilled worker, and we will discuss its construction below. Since probabilities always sum to 1, the last equation in the system in actually redundant. Notice that we can incorporate information about the aggregate state and collect terms from the first row of A.24 to get

$$\mathbf{\Omega_t} = \begin{bmatrix} \xi_t \\ \pi_{\mathbf{t}} \end{bmatrix} = \begin{bmatrix} \chi_t \cdot \kappa \\ s_h \cdot \mathbf{\Phi_{12}}_{,\chi_{\mathbf{t}}} \end{bmatrix} + \begin{bmatrix} (1-g\Delta) & \mathbf{0}'_{\mathbf{dim}(\pi)} \\ \mathbf{0}_{\mathbf{dim}(\pi)} & \mathbf{\Phi_{11}}_{,\chi_{\mathbf{t}}} - s_h \mathbf{\Phi_{12}}_{,\chi_{\mathbf{t}}} \mathbf{1}'_{\mathbf{dim}(\pi)} \end{bmatrix} \mathbf{\Omega_{t-\Delta}}, \tag{A.25}$$

which gives direct expressions for c_{ξ_t} and A_{χ_t}. The probability that $\xi_t = 1$ is $\omega\Delta$, our discrete approximation of the Poisson arrival rate.

Given the MS-VAR representation above, we can leverage closed form results from Bianchi (2016), who demonstrates how to find the ergodic mean and variance-covariance matrix of the stationary distribution for a general class of MS-VARs from which ours is special case. This allows us to compute in closed-form the average level of $E[\xi_t]$ (directly), $E[H_t]$ (by aggregating over the stationary distribution $E[\pi_t]$), and $E[L_t]$ (via the identity $E[L_t] = 1 - E[H_t]$), respectively.

Once we've solved for the ergodic steady state, we can begin calculating the model moments which

[22] Notice that the vector of probabilities $\pi_{\mathbf{t}}$ omits two probabilities. First, we drop the fraction of workers in the absorbing $\theta = 0$ state since it is fixed at $1 - s_h$. Second, we also omit the fraction of workers at the last grid point, because the amount of mass there is implied by the restriction that probabilities sum to one.

correspond to our empirical calibration targets. We begin with our targets related to aggregate quantities. To construct impulse responses in Figure 5, we compute expected paths of H_t, L_t, and ξ_t when there is a single breakthrough event at time t and where we assume that $\mathbf{\Omega_{t-\Delta}} = E[\mathbf{\Omega_t}]$. We can construct an analogous path for the variables if no breakthrough occurs at t. We then compute the corresponding levels of output, wages, and the labor share along these paths. Finally, to translate into standard deviation units, we multiply by an annualization factor $\sqrt{\Delta\omega(1-\omega)}$—that is the volatility of the number of breakthroughs in a year—for aggregate moments. For calibration purposes, we target the impact responses of output and the labor share.

Our remaining calibration targets are computed using the cross-sectional distribution of earnings implied by the model. As with aggregate impulse responses, we compute these cross-sectional moments local to the mean of the stationary distribution of $\mathbf{\Omega_t}$. Likewise, we allocate mass from the stationary distribution of earnings to five income bins defined using the same earnings rank breakpoints defined in our empirical analysis. We then compute the closest model analogue to our empirical estimates by deriving analytical expressions for coefficients of a regression of $\mathbf{y_t}$, the model-implied log change in earnings (or an indicator for the probability of being in the bottom 10% of income declines in a given period), on model observables $\mathbf{x_t}$, including income bin group dummies; a χ_t indicator (analogous to our industry-time FE); and, $\chi_t \times d_{i,t}$ (a proxy for technology exposure) interacted with income group dummies. We do so by computing analytical expressions for $E[\mathbf{x_t x_t'}]$ and $E[\mathbf{x_t y_t}]$, which allows us to compute OLS regression coefficients of interest in closed form.[23] We then multiply coefficients on exposure by $\sqrt{\Delta\omega\alpha(1-\omega\alpha)}$, which equals the volatility of a sum of the high frequency exposures over the course of a year.

Model Calibration

Here we outline details regarding the model calibration. Table 12 lists the model parameters; the final 11 parameters in the table are targeted by the model calibration, while the first 4 are pre-calibrated (see main text for explanation of our choices for pre-calibrated parameters). Table 13 lists the 13 target moments in the model calibration. We choose the parameter vector Θ by minimizing the distance between the output of the model $\hat{X}(\Theta)$ and the data X,

$$\hat{\Theta} = \arg\min_{\Theta} \left(X - \hat{X}(\Theta)\right)' W \left(X - \hat{X}(\Theta)\right). \tag{A.26}$$

The elements of the vector $X - \hat{X}(\Theta)$ in (A.26) are scaled by the absolute value of the empirical moments in order to express model moment deviations in terms of percent differences relative to their counterparts in the data. Since the model is overidentified, our choice of weight matrix W is important for achieving a balanced fit across the set of target moments. We use a diagonal matrix for W with the following weights along the diagonal:

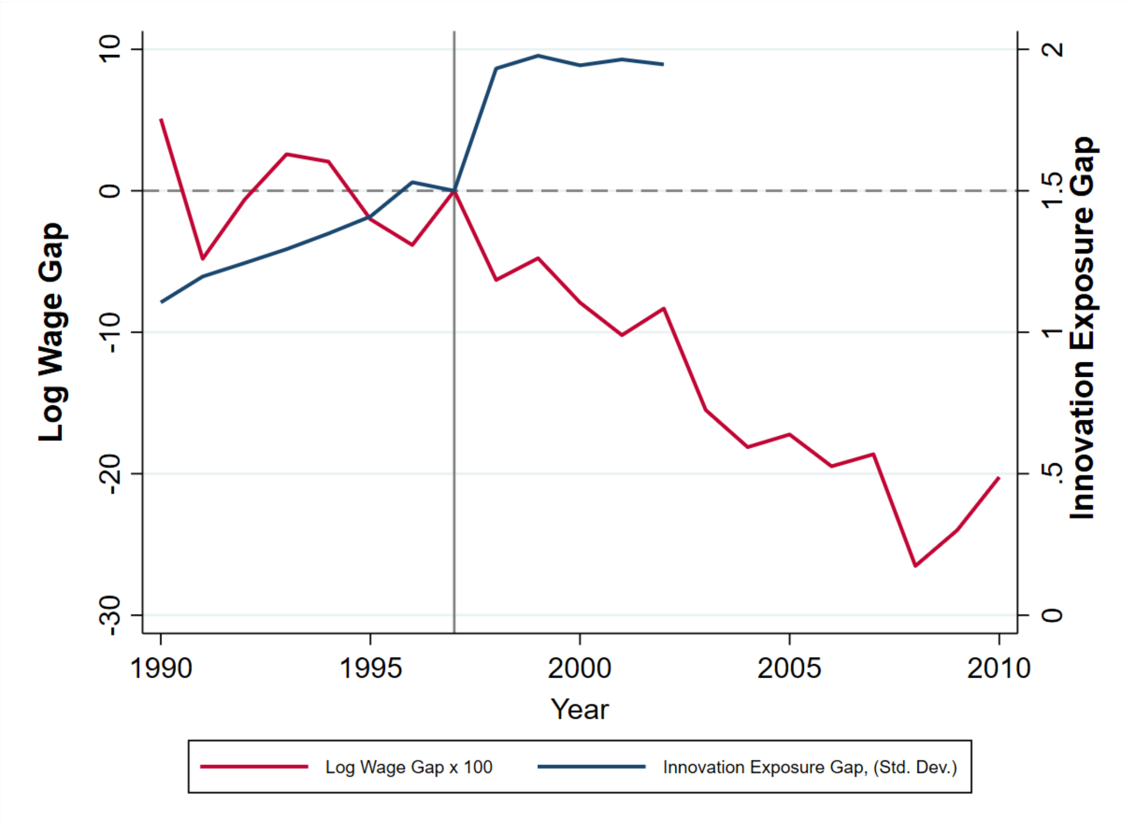
[23] While these expressions are straightforward to derive, they do not add much intuition so we suppress them here.

Moment	Weight
Labor share, response to ξ	0.50
Skill premium (p75 / p25 ratio), average	0.75
Labor productivity, response to ξ	1.0
Worker earnings growth response to ξ	
0 to 25-th percentile	0.25
25 to 50-th percentile	0.1875
50 to 75-th percentile	0.1875
75 to 95-th percentile	0.25
95 to 100-th percentile	0.5625
Difference between 95 to 100-th and 75 to 95-th	0.50
Likelihood of large wage declines in response to ξ	
0 to 25-th percentile	0.625
25 to 50-th percentile	0.625
50 to 75-th percentile	1.0
75 to 95-th percentile	1.0
95 to 100-th percentile	0.625
Difference between 95 to 100-th and 0 to 25-th	0.125

Our weighting matrix generally emphasizes target moments that are on the margin harder to hit. These include the industry-level labor productivity response to changes in technology and the worker level left tail income responses (particularly in the middle of the earnings distribution). In order to ensure that the model generates a more accurate slope in technology-induced drops in average earnings growth and left-tail income outcomes across the income distribution, we also add to X two moments that are linear functions of other moments. These are the gap between the top and second-highest income bin for earnings growth responses and the gap between the top and lowest bin for left tail marginal effects. As shown in Table 13, the resulting calibration provides a fit that is close to the data quantitatively for all the target moments, while also capturing all the key qualitative patterns in the data.

Appendix Figures and Tables

Figure A.1: Example: Order Clerks versus All Other Clerk Occupations



Note: 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 our CPS Merged Outgoing Rotation Groups sample.

Figure A.2: Dispersion in worker earnings: decomposition

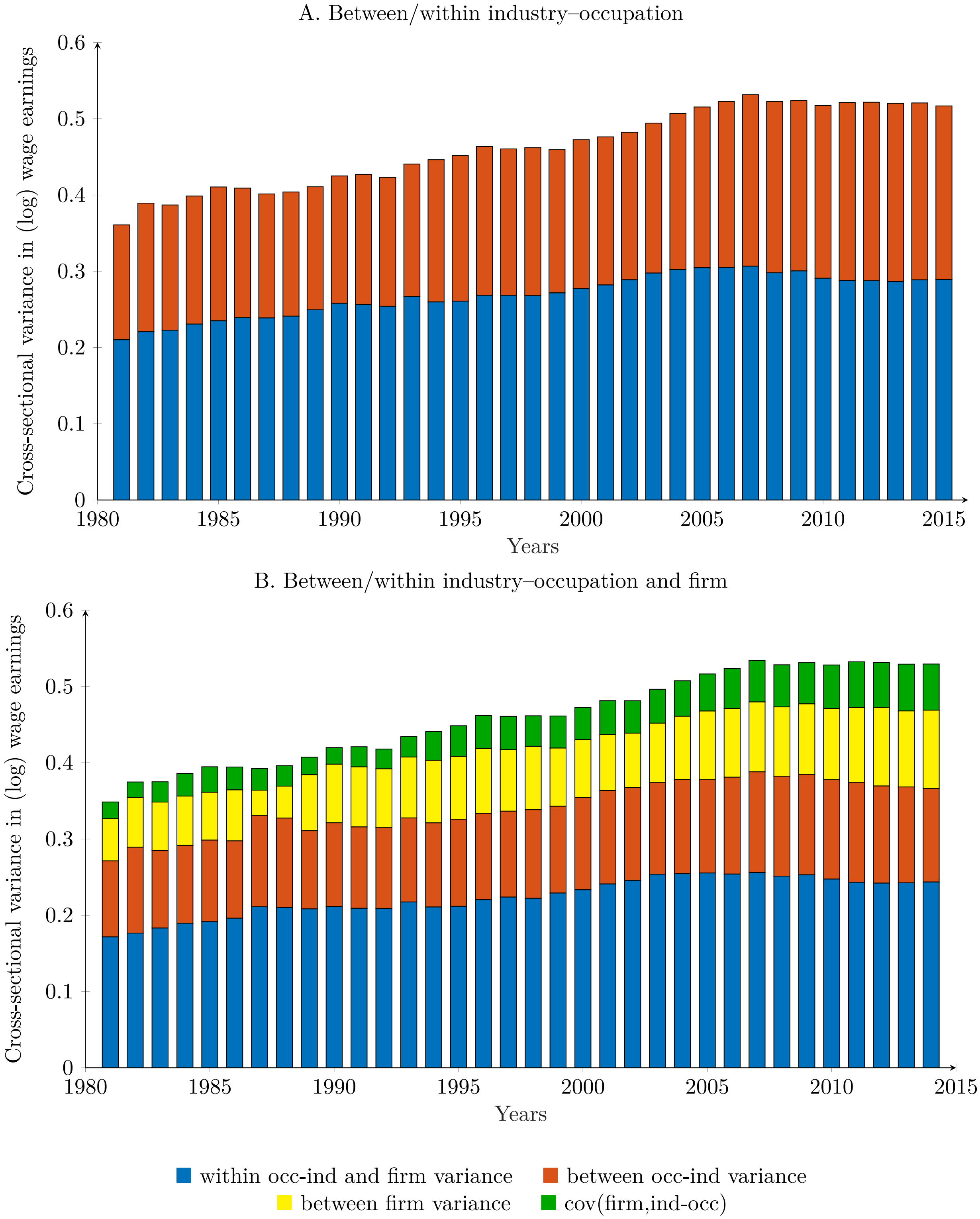
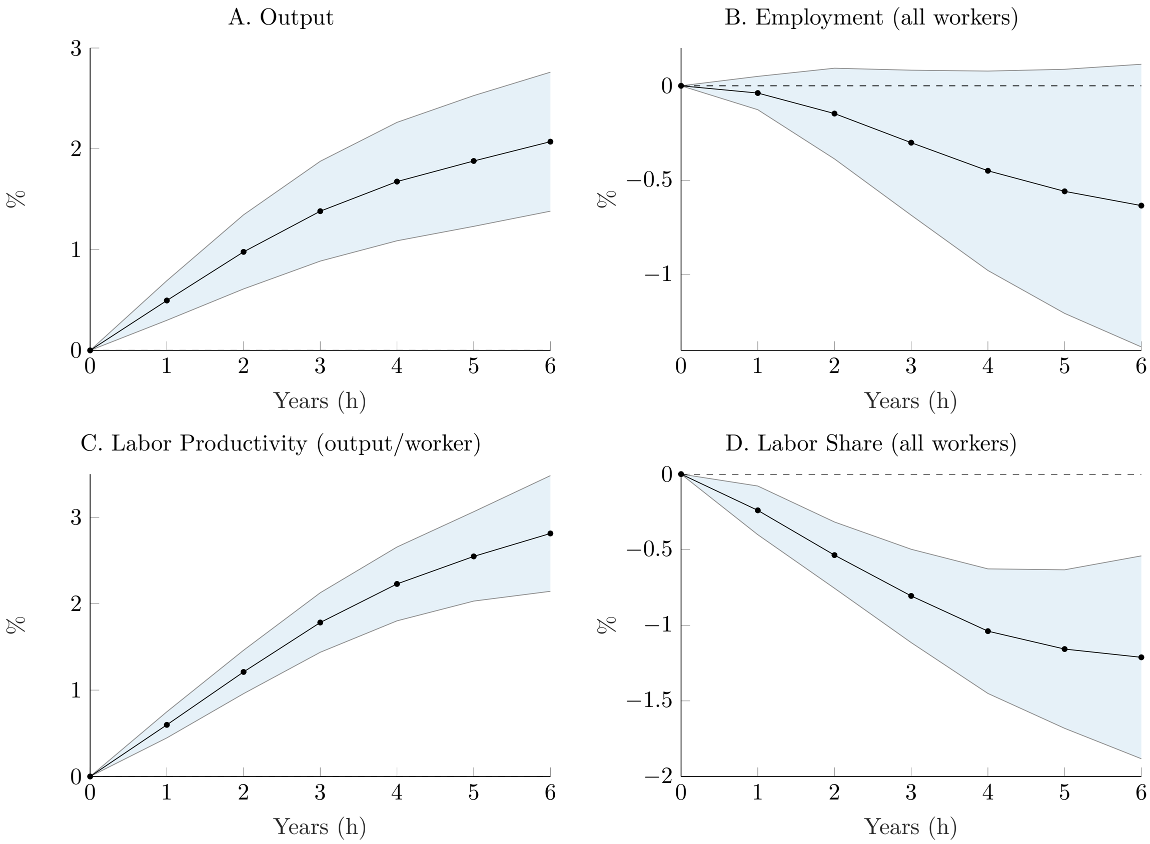


Figure graphs the share of variance in log earnings into within- and between- occupation-industry components (panel A); and within- and between- occupation-industry and firm-type components (panel B) for each year in our sample. See Appendix A.7 for details.

Figure A.3: Innovation: Productivity vs Labor Share



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

$$\log X_{I,t+h} - \log X_{I,t} = \alpha(h) + \beta(h)\,\psi_{I,t} + c(h)\,\mathbf{Z}_{I,t} + \varepsilon_{I,t}, \qquad h = 1\ldots 6 \text{ years.}$$

The main independent variable $\psi_{I,t}$ is an index of innovation in industry I in year t, constructed as follows. First, we assign breakthrough patents to industries using the probabilistic patent CPC tech class to industry crosswalk from Goldschlag et al. (2020). We scale $\psi_{I,t}$ by US population and normalize to unit standard deviation. Controls $Z_{I,t}$ include year fixed effects and lagged 5-year growth rate of the dependent variable. Standard errors are clustered by industry, and corresponding 90% confidence intervals are shaded.

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: Breakthrough patents most related to tasks performed by order-fulfillment clerks

US. Pat. #	Similarity (ρ)	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.3: 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
	(0.43)	(0.58)				(0.63)	(0.72)
Routine Task Intensity Rank			-2.16			-2.31	-1.64
			(0.44)			(0.61)	(0.54)
Robot Exposure Rank				-1.69		0.52	-0.15
				(0.43)		(0.71)	(0.62)
Software Exposure Rank					-1.20	-0.34	0.050
					(0.53)	(0.73)	(0.70)
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
	(0.45)	(0.57)				(0.63)	(0.73)
Routine Task Intensity Rank			-2.60			-2.48	-1.85
			(0.41)			(0.54)	(0.46)
Robot Exposure Rank				-2.27		-0.13	-0.70
				(0.45)		(0.64)	(0.56)
Software Exposure Rank					-1.43	0.057	0.32
					(0.54)	(0.68)	(0.65)
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: Table shows results from estimating

$$\frac{100}{h}\Big(\log Y_{i,j,2012} - \log Y_{i,j,1980}\Big) = \alpha + \alpha_j + \beta\eta^{\text{Pctile}}_{i,1980} + \delta X^{\text{Pctile}}_i + \epsilon_{i,j}$$

Here i indexes occupations and j indexes industries; we report results for $h = 32$ years using data from 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. 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 (2020), depending on the specification. Because the Webb (2020) 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. We report standard errors clustered at the occupation level in parentheses beneath coefficient estimates. Coefficients are multiplied by 100 for readability, and observations are weighted by employment share in 1980.

Table A.4: 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: 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.5: Technology exposure and worker earnings growth by income rank, alternative income rankings

	Income Ranking Method					
					Residual Earnings	
	Baseline	Drop Recent Hires	2-yr Avg Income	Firm Adjusted	Age/Gender	Union
	(1)	(2)	(3)	(4)	(5)	(6)
0–25th percentile	-1.85	-1.95	-2.07	-1.94	-2.03	-2.01
	(0.49)	(0.51)	(0.46)	(0.50)	(0.51)	(0.80)
25-50th percentile	-1.49	-1.20	-1.44	-1.43	-1.27	-1.80
	(0.43)	(0.43)	(0.37)	(0.44)	(0.44)	(0.66)
50-75th percentile	-1.85	-1.56	-1.74	-1.65	-1.80	-1.80
	(0.39)	(0.40)	(0.37)	(0.43)	(0.43)	(0.68)
75-95th percentile	-2.52	-2.11	-2.27	-2.10	-2.17	-2.20
	(0.42)	(0.41)	(0.39)	(0.42)	(0.42)	(0.71)
95–Top	-4.21	-3.70	-3.66	-3.66	-4.43	-4.48
	(0.58)	(0.56)	(0.58)	(0.68)	(0.61)	(1.11)
95–Top vs 25–95th percentile	-2.26	-2.08	-1.84	-1.94	-2.63	-2.59
(p-val,%)	0.00	0.00	0.01	0.04	0.00	0.19
95–Top vs 0–25th percentile	-2.36	-1.75	-1.59	-1.72	-2.31	-2.41
(p-val,%)	0.10	1.15	2.64	0.58	0.02	0.63
0-25th percentile vs 25–75th percentile	0.10	-0.33	-0.25	-0.22	-0.28	-0.15
(p-val,%)	77.70	34.78	50.74	46.26	40.71	72.75

Note: Table shows the estimated slope coefficients β (times 100) from equation (8) 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 (6) use different methods for ranking workers based on earnings. Column (1) represents our baseline method of sorting workers each year within occupation–industry. Column (2) uses our baseline income sort but drops workers who were hired within the last year. Column (3) ranks workers on average earnings over the past two years. Column (4) adjusts workers' earnings for average firm wages by subtracting off the log average wage of a worker's employer in the Longitudinal Business Database from the worker's log wage. Column (5) residualizes log earnings with respect to yearly fixed effects for occupation–industry; commuting zone; and 10-year age bin (25-35, 35-45, 45-55) interacted with gender. Column (6) restricts to subsample of workers who are in-universe for the CPS union membership question, and residualizes log earnings with respect to yearly fixed effects for occupation, industry, worker union status, and commuting zone. All specifications include industry $\times$ year, occupation $\times$ year, baseline within occupation–industry income bin $\times$ year fixed effects, and dummies for the income bin version used in each specification. The bottom panel of the table reports differences within each column between the top and bottom income bin coefficients; top income bin and average of the middle three income bins; and, bottom income bin and average of the middle three income bins, along with the corresponding p-values (in percent units) for a two-sided test of coefficient equality.

Table A.6: Technology exposure and worker earnings growth by income rank, comparison across unionized vs non-unionized workers

Worker earnings rank (rel. to occ × ind group)	Industry Unionization		Worker In Union	
	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: Table shows the estimated slope coefficients β (times 100) from equation (8) 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 in parentheses beneath coefficient estimates, and normalize $\xi_{i,t}$ to unit standard deviation. In the first two columns 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. The bottom panel of the table reports differences within each column between the top and bottom income bin coefficients; top income bin and average of the middle three income bins; and, bottom income bin and average of the middle three income bins, along with the corresponding p-values (in percent units) for a two-sided test of coefficient equality.

Table A.7: Technology exposure and worker earnings risk

Worker earnings rank (rel. to occ × ind)	All Workers	Stay in Firm: Same EIN for at least next 5 yrs
0–25th percentile	0.82	0.02
	(0.22)	(0.09)
25–50th percentile	0.61	0.13
	(0.21)	(0.10)
50–75th percentile	0.68	0.17
	(0.21)	(0.10)
75–95th percentile	0.91	0.19
	(0.21)	(0.10)
95–Top	1.62	0.36
	(0.29)	(0.14)
95–Top vs 25–95th percentile	0.89	0.20
(p-val, %)	0.01	10.08
95–Top vs 0–25th percentile	0.81	0.35
(p-val, %)	0.45	1.22
0–25th percentile vs 25–95th percentile	0.09	-0.14
(p-val, %)	52.13	2.02

Note: Table shows the estimated slope coefficients β (times 100) from a version of equation (8) 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 in parentheses 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. The bottom panel of the table reports differences within each column between the top and bottom income bin coefficients, the top income bin and average of the middle three income bins, and bottom income bin and average of the middle three income bins, along with corresponding p-values (in percent units) for a two-sided test of coefficient equality.