New Frontiers: The Origins and Content of New Work, 1940–2018*

David Autor
MIT & NBER

Anna Salomons
Utrecht University & IZA

Bryan Seegmiller
MIT Sloan

July 26, 2021

Abstract

Recent theory stresses the role of new job types ('new work') in counterbalancing the erosive effect of task-displacing automation on labor demand. Drawing on a novel inventory of eight decades of new job titles linked to United States Census microdata, we estimate that the majority of contemporary employment is found in new job tasks added since 1940 but that the locus of new task creation has shifted—from middle-paid production and clerical occupations in the first four post-WWII decades, to high-paid professional and, secondarily, low-paid services since 1980. We hypothesize that new tasks emerge in occupations where new innovations complement their outputs ('augmentation') or market size expands, while conversely, employment contracts in occupations where innovations substitute for labor inputs ('automation') or market size contracts. Leveraging proxies for output-augmenting and task-automating innovations built from a century of patent data and harnessing occupational demand shifts stemming from trade and demographic shocks, we show that new occupational tasks emerge in response to both positive demand shifts and augmenting innovations, but not in response to negative demand shifts or automation innovations. We document that the flow of both augmentation and automation innovations is positively correlated across occupations, yet these two faces of innovation have strongly countervailing relationships with occupational labor demand.

Keywords: Technological Change, New Tasks, Augmentation, Automation, Demand Shifts
JEL: E24, J11, J23, J24

*We thank Daron Acemoglu, David Deming, Maarten Goos, Gordon Hanson, Lynda Laughlin, the expert staff of the U.S. Census Bureau, and numerous seminar and conference participants for insightful comments that have improved the paper; and Fiona Chen, Caroline Chin, Grace Chuan, Rebecca Jackson, Zhe Fredric Kong, Edwin Song, Ishaana Talesara, Liang Sunny Tan, Jose Velarde, and Whitney Zhang for expert research assistance. We thank Dimitris Papanikoloau for sharing a database of raw patent texts; and Sammy Mark, Fredrick Soo, and Mary Wakarindy from Digital Divide Data for hand-keying historical Census Alphabetical Index records. We gratefully acknowledge support from the Washington Center for Equitable Growth, the MIT Work of the Future Taskforce, The Carnegie Corporation, Schmidt Futures, the Smith Richardson Foundation, Accenture, and Instituut Gak.

1 Introduction

A burgeoning economic literature analyzes how rapidly evolving technologies—information and communications technologies, artificial intelligence, robotics—affect employment, skill demands, and earnings levels. Focusing on the substitution of machines for workers in tasks where automation has rising comparative advantage, this work anticipates and interprets the decline of middle-skill employment in high income countries (aka, job polarization), documents the concentrated impact of industrial robotics on labor demand in heavy manufacturing industries and in manufacturing-intensive communities, and explores how artificial intelligence may change the structure of occupations.[1]

This body of work is however comparatively silent—with key exceptions, discussed below—on the flip side of the ledger: the augmentation of human labor and the generation of new work activities that demand this labor. Indeed, research on the impact of technological change on employment has primarily treated the set of human job tasks as finite and static, meaning that as automation proceeds, labor is slowly shunted into an ever-narrowing scope of activities, as in Susskind (2020). But casual observation and historical evidence suggest the opposite: even as employment in labor-intensive sectors has eroded—in agriculture, textiles, and mining—the scope and variety of labor-demanding activities has arguably expanded, e.g., in finance, medicine, software, electronics, healthcare, entertainment, recreation, personal care, and many other domains—a phenomenon that Acemoglu and Restrepo (2019) refer to as labor reinstatement.

Though no economic historian would deny the importance of new work creation, formal analysis of this topic has likely lagged because technology-labor complementarity is a residual force in the widely-applied 'task framework' (Acemoglu and Autor, 2011). Tasks that are not substituted are implicitly complemented—and hence this formulation provides little conceptual or empirical guidance for exploring this complementarity in practice. Recent work by Acemoglu and Restrepo overcomes this theoretical limitation (Acemoglu and Restrepo, 2018, 2019), but an empirical challenge remains. While it is relatively straight-

[1]On job polarization, see Autor et al. (2003), Autor et al. (2006), Goos and Manning (2007), Goos et al. (2009), Acemoglu and Autor (2011), Autor and Dorn (2013), Goos et al. (2014) Michaels et al. (2014), Akerman et al. (2015), Arntz et al. (2017), Bárány and Siegel (2018), Dillender and Forsythe (2019), Cortes et al. (2020), and Harrigan et al. (2021). On industrial robotics, see Graetz and Michaels (2018), Chiacchio et al. (2018), Humlum (2019), Acemoglu et al. (2020b); Acemoglu and Restrepo (2020), De Vries et al. (2020), Bonfiglioli et al. (2020), Faber (2020), and Dauth et al. (2021). On the potential impact of artificial intelligence on jobs, see Brynjolfsson et al. (2018), Felten et al. (2019, 2018), Acemoglu et al. (2020a), Alekseeva et al. (2020), Babina et al. (2020), Grennan and Michaely (2020), and Webb (2020).

forward to quantify the set of job tasks and encompassing occupations that are substituted by automation, there is almost no direct measurement of the emergence of new work tasks within occupations and industries or over time.

This paper systematically studies the emergence of new work in the United States between 1940 and 2018, building on path-breaking empirical work by Lin (2011) and theoretical work by Acemoglu and Restrepo (2018). Our objectives are to consistently measure the evolution of new work over eight decades, document its changing locus and relationship to the occupational structure of employment, and explore the forces that explain where new work appears and where old work disappears.

We formalize our main hypotheses in the context of a stylized two-sector task model, building upon Acemoglu and Restrepo (2018), Acemoglu and Autor (2011) and Autor et al. (2003), which draws economic linkages between new task creation, task automation, incentives for innovation, and the locus and attendant skill demands of new work. We posit that new job tasks derive from two primary sources. The first is augmentation, meaning the creation of new production processes (e.g., the Pilkington float glass process, the semiconductor fabrication system), new technologies (e.g., the internal combustion engine, the Global Positioning System), and entirely new products or industries (e.g., commercial aircraft, photovoltaic solar collectors). These sources of innovation create new demands for expert knowledge and specific competencies that drive occupational specialization and hence the creation of new work tasks. The second source of new task creation is changes in market size—stemming for example from trade, demographic shift, immigration, etc.—that increase or depress the value of occupational outputs. Even absent specific technological advances, we conjecture that positive demand shocks catalyze specialization and differentiation of the goods or services produced by an occupation, again spurring new task creation.

Following Acemoglu and Restrepo (2018), we endogenize task creation and task automation in the model by allowing demand shifts to raise the value of occupational outputs, thus generating incentives for entrepreneurs to introduce innovations, which demand new tasks. In this setup, positive demand shocks spur new task creation within occupations while negative demand shocks slow new task creation. We contrast these new task creating forces theoretically and empirically with task automation, which allows for the reallocation of existing tasks from labor to capital. While positive occupational demand shifts also spur innovations that automate tasks, the new task creation margin dominates in our framework. A key final prediction from our conceptual framework is that task-creating and task-displacing forces have opposing labor demand effects: occupational employment and wagebills unambiguously expand with new task creation and contract with task automation.

We develop two unique data sources to test the implications of our framework. Building on pioneering work by Lin (2011), we first construct a database of new job tasks introduced during the period of 1940 through 2018. This database is sourced from nearly a century of internal reference volumes developed and used by U.S. Census Bureau employees to classify the free-text job descriptions supplied by Census respondents into occupation and industry categories in each decade. While Census tabulations and public use data sources report several hundred distinct occupation and industry codes in each Census year (which we call 'macro-titles'), these titles reflect concatenations of approximately 30,000 occupational and 20,000 industry-level 'micro-titles' enumerated in the *Census Alphabetical Index of Occupations and Industries* in each decade between 1930 and 2018 (US Census Bureau). Critically, these indexes are updated during the processing of each decade's Census to reflect new write-in titles detected by Census coders. By comparing successive editions of the Census Alphabetical indexes, we are therefore able to track the emergence of new micro-titles across decades. For example, the micro-title of "Technician, fingernail" was added to the Census Alphabetical Index in 2000, and the micro-title of "Solar photovoltaic electrician" was added in 2018.[2]

The second novel data source is a comprehensive, quantitative classification of the flow of patents over nine decades that identifies innovations which, on the one hand, complement occupational outputs and, on the other, substitute for labor-using occupational inputs. We construct these innovation measures using natural language processing (NLP) tools to map the content of U.S. utility patents to the domain of occupations between 1930 and 2018. Following Kogan et al. (2019), we represent documents as on weighted averages word embeddings, which are geometric representations of word meanings, to measure the distance between patent texts and occupational descriptions.[3]

With these tools, we develop two conceptually distinct measures of innovation flows. The first captures innovations that may complement the output of occupations, potentially creating new demands for occupational expertise or occupational services. We construct this

[2]Our work extends and expands upon Lin (2011), who constructed a new work inventory over 1980–2000 based on comparisons of Dictionary of Occupational Titles (DOT) records from 1965, 1977, and 1991, and Census Alphabetical Index of Occupations data from 1990 and 2000.

[3]Relative to conventional measures of text similarity (for example, the commonly used bag of words approached outlined by Gentzkow et al. (2019)), the key advantage of word embeddings is that they account for synonyms—which is crucial in our context since patent texts and occupational descriptions may use different terminology for similar concepts.

index by calculating the overlap between patent texts and the micro-titles from the Census Alphabetical Index associated with each industry and occupation to identify innovations that are aligned with occupational outputs. For example, in 1998, the U.S. Patent and Trademark Office (USPTO) granted patent US5924427A for a "Method of strengthening and repairing fingernails". Our algorithm links this patent to the Census macro-occupation of "Miscellaneous personal appearance workers," which encompasses the micro-title of "Technician, fingernail" which entered the Census Alphabetical Index in the year 2000. Similarly, our algorithm links the 2014 patent US7605498B2 "Systems for highly efficient solar power conversion" to the macro-industry of "Electronic component and product manufacturing, not elsewhere classified". In turn, this industry links to the micro-occupational title of "Solar photovoltaic electrician", which entered the Census Alphabetical Index in the year 2018.

Our second patent-based measure captures innovations that may substitute for the inputs of occupations. For this index, we follow Kogan et al. (2019) and Webb (2020) in using NLP tools to identify the overlap between the content of patents and the tasks that workers perform, as described by the Dictionary of Occupational Titles (U. S. Department of Labor, Employment and Training Administration, 1991).[4] For example, in 1977, the USPTO granted patent US4141082A for a "Wash-and-wear coat". Our algorithm links this patent to the macro-occupation of "Laundry and dry cleaning workers." Similarly, our algorithm links 1976 patent US3938435A, "Automatic mail processing apparatus", to the macro-occupation of "Mail and paper handlers."

Our empirical findings are as follows. First, the majority of contemporary employment is found in new job tasks added since 1940, and the changing locus of innovative activity predicts where new tasks emerge. Further, augmentation and automation innovations have distinct relationships to new task emergence. Augmentation innovations strongly predict the locus of new task emergence, as measured by job titles, across occupations and over time. Automation innovations do not predict new title emergence, despite their positive correlation with augmentation innovations at the level of occupations.

Alongside augmentation innovations, exogenous demand shifts for occupational outputs also predict when and where new work emerges. Increases in demand favoring a sector lead to the emergence of new work tasks, while conversely, adverse demand shocks retard the emergence of new work tasks in exposed sectors. These shifts help account for the emergence

[4]In related work, Felten et al. (2019, 2018) and Brynjolfsson et al. (2018) develop measures of the exposure of occupations to advances in artificial intelligence and machine learning.

of new tasks over the last four decades in lower-paid personal services, an occupational category which has been relatively unexposed to innovative activity but where new work has nonetheless emerged.

Lastly, we find robust evidence that labor demand shifts outward with new task creation and inward with task automation: employment and wagebills expand in occupations exposed to augmentation innovations and contract in occupations exposed to automation innovations. Bearing in mind that augmentation and automation innovations are positively correlated, these countervailing relationships with occupational labor demand are all the more striking.

This work contributes to three economic literatures. A first studies the interplay between supply, demand, technologies, and institutions in shaping the long-run evolution of skill demands, occupational structure, and wage inequality (Goldin and Margo 1992, Katz and Murphy 1992, Katz and Autor 1999, DiNardo et al. 1996, Acemoglu 1998, Autor et al. 1998, Card and Lemieux 2001, Goldin and Katz 2008, Autor et al. 2020, Haanwinckel 2020). A foundational assumption of this literature is that technological change shapes the skill bias of labor demand. Our paper adds nuance and specificity to this idea by linking changes in the structure of occupational demands to the shifting locus of innovation over eight decades. We show that new work is a quantitatively large contributor to aggregate employment change, that it emerges where innovative activity is focused, and that the focus of this activity has shifted from middle-educated, production-oriented sectors, such as mining, manufacturing processes, and transportation, to primarily highly-educated sectors, including electricity and electronics and instruments and information.

Our paper also speaks to a contemporary literature exploring how automation technologies substitute for existing work, as measured by occupational structure or job tasks (see citations in footnote 1). We build on Kogan et al. (2019); Mann and Püttmann (2020); Webb (2020), who devise textual analysis methods to identify innovations recorded in patents that potentially overlap the tasks performed by occupations, as well as papers by Brynjolfsson and Mitchell (2017); Brynjolfsson et al. (2018); Felten et al. (2019, 2018) that predict which occupational tasks can be performed by artificial intelligence. Distinct from this literature, we develop and empirically verify a method to identify innovations that generate *new* work tasks by complementing occupational outputs. Our finding that employment rises in occupations exposed to augmentation innovations is novel to the literature and illustrates the power of distinguishing among innovations according to their economic content.

Our paper contributes most directly to research assessing the micro- and macroeconomic origins of new work, including Goldin and Katz (1998), Lin (2011), Acemoglu and Restrepo

(2018), Acemoglu and Restrepo (2019), Atack et al. (2019), Frey (2019), Atalay et al. (2020), and Deming and Noray (2020). Conceptually, our approach builds upon Acemoglu and Restrepo (2018), who model the interplay between automation and new task creation in shaping labor demand, productivity, and factor shares. Following their approach, we treat task creation and task automation as endogenous: Entrepreneurs supply innovations to either complement workers (task creation) or complement capital (task automation) in response to factor prices. We additionally allow for sectoral demand shifts that influence where entrepreneurs find it most profitable to innovate, thus altering the set of occupations in which new tasks emerge.

Lastly, we make empirical contributions to literatures measuring new work and innovation. In measuring new work creation, we extend the ideas pioneered in Lin (2011), while expanding the scope of the analysis from two to eight decades. Our approach is related to but distinct from Acemoglu and Restrepo (2019), who develop a set of ingenious proxies for the appearance of new work based on changes in labor share and in the mix of 3-digit occupations (what we call 'macro-occupations') within industries. In related work, Atalay et al. (2020); Deming and Noray (2020), measure the appearance of new work by analyzing the text of job advertisements. Our paper complements these approaches by providing direct, representative, and time-consistent measurement of new task creation across eight decades. We leverage these task measures by linking both new and existing occupational titles to innovations recorded in patents, building on recent work by Kogan et al. (2019); Mann and Püttmann (2020); Webb (2020). Distinct from this work—and all prior work to our knowledge—we identify innovations that complement occupational outputs, which we hypothesize (and empirically confirm) spur new task creation. Most notably, our analysis successfully draws on a unified corpus of patents to distinguish between innovations that substitute for occupational inputs and those that complement occupational outputs. This tool, which may be of independent interest, permits us to analyze new task creation and task automation concurrently, and to show that these forces are positively correlated at the occupational level and yet have countervailing relationships with occupational employment growth.[5]

[5]We also build on the vast literature, originating with Griliches (1981); Hall et al. (2001); Jaffe et al. (1993), that uses patents to study knowledge spillovers, innovation networks, the value of innovation and its relationship to rent creation, public-private R&D complementarities, and innovation responses to taxation, among many other topics. See Hall and Harhoff (2012); Moser (2016) for recent reviews of (aspects of) this literature.

The paper proceeds as follows. Section 2 details our methods for identifying new work and for linking task-replacing and task-creating innovations to occupations. Section 3 provides descriptive evidence on the locus of new work over 1940–2018. Section 4 develops an illustrative theoretical model that motivates and guides our empirical work. Section 5 tests two foundational assumptions of our model: that output-complementary innovations are associated with new work emergence; and that input-substituting innovations are *not* associated with new work emergence. Section 6 examines whether new task creation responds elastically to negative demand shocks stemming from globalization, and to positive demand shifts stemming from demographic changes. The final empirical component of the paper, Section 7, assesses whether augmentation and automation innovations have distinct, countervailing relationships with occupational employment and wagebill growth, as our model implies. Section 8 concludes.

2 Data and Measurement

We start by constructing a novel and detailed inventory of new job titles, spanning 1940–2018, which we link to representative worker data from the U.S. Census and American Community Survey. To test our hypotheses, we further construct measures of occupations' exposure to both augmentation and automation innovations and to demand shifts.

2.1 Measuring new work

Our work leverages Census historical coding volumes for occupations for the years 1930 through 2018 (the Census Alphabetical Index of Occupations) and for industries (the Census Alphabetical Index of Industries) for 1940 through 2018. For brevity, we refer to these volumes as the CAI hereafter. Each Census coding volume contains around 35,000 occupation and 15,000 industry 'micro' titles in each year, each classified to a more aggregated ('macro') Census occupation or industry code. These catalogues of micro titles serve as reference documents for Census coders, who classify individual Census write-ins for job title and industry of employment.[6] This process has been performed consistently since 1900 and is

[6] In the post-2000 years of our analysis, we use data from the American Community Survey (ACS) rather than the U.S. Census itself. The ACS is the successor to the Census long form and retains the industry and occupation coding and processing structure of earlier Census files. The CAI volumes are regularly updated for processing of the ACS.

illustrated for occupations in the American Community Survey (ACS) in Figure 1.

When Census coders encounter written-in occupational titles that cannot be linked to any existing micro title, the Census bureau, after internal review, adds a new micro title to that decade's coding volume. For example, the micro occupation title "Artificial Intelligence Specialist" was added in 2000, since a sufficient number of Census respondents reported this novel title (or something similar to it) as their main occupation. This new micro title was then classified to the broader Census ('macro') occupational title "Computer Scientists and Systems Analysts", which appears in published Census tabulations and public use data sets.[7] Although there is not a specific numeric threshold for the number of times that a write-in title must recur to trigger the addition of a new micro-title, it is apparent that titles are added only after they draw notice from Census coders. For the purpose of measuring new work, it is desirable that these new titles are not so esoteric as to be insignificant in the U.S. working population. This explains why Mental-Health Counselor is a new title in 1970, Artificial Intelligence Specialist is new in 2000, and Sommelier is new in 2010: Although there were surely workers performing these job types in earlier decades, the particular specialization was too rare to warrant inclusion beyond a generic counselor, computer science, or waitstaff title. Comparing successive editions of the CAI occupation and industry coding volumes across decades thus allows us to uncover the newly added jobs and activities that are becoming prevalent enough to register.

The Census does not highlight or separately list newly added titles, and coding volumes are also revised for other reasons, such as renaming outdated job descriptions, adding differently phrased variants of the same title, or removing gendered forms.[8] To extract new titles, we compare title lists across decades using fuzzy matching combined with extensive manual revision of 'candidate-new' titles, discarding false positives that emerge from for example rewording, reformatting of the index, or other newly added titles which do not reflect a discernible modification to their preexisting counterparts.[9] Our overarching aim is to retain newly added titles which reflect truly new work, meaning that they add a particular

---

[7] The assignment of micro titles is an intermediate step in the assignment of macro titles to Census records. While each respondent's macro title is permanently attached to her Census record, her intermediate micro title is never recorded.

[8] The Census does not systematically remove extinct titles from the index, as these titles can still help classify write-ins for titles that have become relatively uncommon in the U.S. working population.

[9] For example, we do not count "Software Applications Developer" as a new micro-occupation because "Software Developer" was already present when it was added.

task specialization, work method or tool, or professional or educational requirement.[10] For example, "Clinical Psychologist" is new in 1950 because it was not present in 1940 and is a specialization of the preexisting title of Psychologist.

To see the cumulative force of this process, consider the 1940 Census occupation (macro-title) of Mechanics and Repairmen—Automobile (occupation 332). This occupation encompassed 84 distinct micro-titles in 1940, ranging alphabetically from Alignment Man: Auto Repair Shop to Windshield Man: Auto Repair Shop. In 2018, 78 years later, there were 134 micro-titles in the closely related (though imperfect) successor to the 1940 occupation, Automotive Service Technicians and Mechanics (Census occupation 7200). Of the approximately 50 titles added in the intervening seven decades, many reflect further specialization of existing activities present in 1940—for example, the number of brake repair micro-occupations expanded from four in 1940 to nine in 2018. Additionally, numerous micro-titles recorded in subsequent decades reflect technologies introduced after 1940: Hybrid Car Mechanic, Fuel Injection Servicer, Remote Control Mirror Installer, and three types of automotive air conditioning specialists.

Table 1 documents the diversity of micro-titles added to the CAI in each decade from 1940 through 2018.[11]. The left-hand column of the table reports titles that, akin to automobile mechanics, specialize around new or evolving technologies: Airplane designers in 1950, Engineers of computer applications in 1970, Circuit layout designers in 1990, and Technicians of wind turbines in 2010. While some new tasks have direct technological origins, others may emerge in response to changing tastes, income levels, and market size. The right-hand column of Table 1 reports emerging occupations that have no obvious technological genesis: Tattooers in 1950, Hypnotherapists in 1980, Conference planners in 1990, and Drama therapists in 2018. Such examples motivate looking beyond exclusively technological forces in analyzing the sources of new work creation, focusing specifically on demand channels.

Because the CAI is a coding aid rather than a survey instrument, the Census Bureau does not record or report the count of respondents within micro-titles. We employ Census of Population and American Community Survey (ACS) public use data files to connect these micro-titles to representative data for the years 1940 through 2018. The Census public use files typically report several hundred distinct occupations ('macro' titles) in each year. The

[10] Appendix A.1 provides details.

[11] New titles emerging in 1940 refer to title added between 1930 and 1940, and similarly for subsequent decades.

average macro-occupation concatenates respondents from approximately one-hundred micro-titles. Thus, we draw a many-to-one linkage between micro-titles in the CAI and 'macro' Census occupations in Census public use data.

From this micro-macro match we construct two measures of new work: (1) the number of new 'micro' titles in each 'macro' Census occupation; and (2) the *new title share* in each occupation. Following Lin (2011), the new title share is defined as the ratio of new micro titles to the total number of micro-titles within a Census occupation. As noted above, the Census data do not allow us to observe which workers within a macro-occupation occupy which micro-titles, and our analysis does not for the most part require such information. (Our primary empirical analysis *estimates* the relationship between new titles, innovation flows, and employment changes.) For illustrative purposes, however, we make two imputations below regarding the characteristics of 'new workers': we impute their demographic characteristics (e.g. education and earnings) as the average of the characteristics of workers in their macro occupation by industry cell[12]; and we impute the share of workers in 'new work' within each macro-occupation as the new title share.[13]

By matching the cumulative flow of new titles over eight decades to Census data, we estimate that the majority of contemporary work is found in new tasks added since 1940, as shown in Figure 2. This figure charts the distribution of employment in 1940 and 2018 in twelve exhaustive, mutually exclusive broad occupational categories ordered from lowest to highest-paying, with Farming and Mining occupations on the left-hand side of the scale and Managerial workers on the right-hand side. In the second set of columns (those for 2018), we further distinguish between 2018 employment found in occupational titles that existed in 1940 versus 2018 employment in occupational titles that were added thereafter (i.e., new tasks).[14] Roughly 63 percent of employment in 2018 is found in jobs that did not exist in

[12]Each Census respondent is classified to a macro-occupation and macro-industry. Using both data points to assign characteristics to respondents increases the specificity of the imputation.

[13]The accuracy of the latter approximation is likely to differ over time-horizons. Plausibly, employment in a micro-title will tend to grow following its introduction, so the new work share may overstate the true fraction of workers in a given micro-title in the decade of its introduction but then may understate its fraction one, two, or three decades later. We have explored the validity of this imputation approach using Census Complete Count data for 1940, which contains both macro-titles and the free text write-in micro-titles supplied by Census respondents. We estimate that the count of workers in new titles is strongly increasing in the new title share—though the slope is below one—and that this relationship is more precise when using ordinal share ranks rather than cardinal shares. Details are given in Appendix A.2.

[14]Employment in titles is estimated by constructing a cumulative new title share—summing the number of new titles added over 1940–2018—and dividing this by the total number titles in the 2018 index adjusted for titles that were removed, separately by broad occupation. Details are reported in Appendix A.2.

1940. Among Professionals—the occupational category that added the most workers during these eight decades—this share was 75 percent. Conversely, less than half (48 percent) of employment in Production occupations in 2018 is found in job categories that were not present in 1940. Notably, Production had the second lowest employment growth out of these broad occupational categories during the past eight decades, the other being Farming.

Our primary analyses focus on the *distribution* of new work added by decade rather than the absolute numbers of new titles added, which also depends on available resources at the U.S. Census Bureau for revising the index. By focusing on the distribution of new titles added in a given decade—representing the flow of new titles between decade t and $t-1$—we require only that efforts to keep the index representative within a decade are not biased towards any particular set of occupations.

2.2 Measuring output-complementary and input-substituting innovations

Our second empirical task is to measure the exposure of occupations to innovations that may complement their outputs or substitute for their inputs. We measure innovation using patent data, following a large literature. Patent data were obtained by Kelly et al. (2020) from the USPTO patent search website for patents issued from 1976–2015, and from Google Patents prior to 1976. We extend the Kelly et al. (2020) sample of patent issued from 2015 to 2018 by scraping the patent text from the Google Patents website. To link these data to their relevant occupations and industries, we use the entire text of each patent.[15]

The locus of innovation has shifted strongly across technological categories over decades, as Figure 3 illustrates by grouping citation-weighted patents into broad technology groups following Kelly et al. (2020).[16] While patenting activity was largely concentrated in Transportation, Manufacturing Processes, and Engineering, Construction, and Mining at the start of the twentieth century, innovations in Chemistry and Metallurgy gained prominence from 1940 onward. Since 1980, innovative activity has shifted strongly towards digital technologies, seen in the growing share of patents in Instruments and Information, and Electricity

[15] After 1976, patent texts are divided into their abstract, description, and claims sections. Prior to that time, the patent documents come in a single block of text.

[16] We also follow Kelly et al. (2020) in using data from Berkes (2018) to get citation counts for patents issued prior to 1976. We compute citation counts from the USPTO PatentsView database (available for download here) for patents issued 1976 and later.

and Electronics. Indeed, from 2010 forward, more than half of all new patents are found in these two classes. This period has additionally seen a marked increase in the importance of patents in health-related technologies. The fact that the locus of innovation has shifted technology classes during these decades implies, as we document next, that different types of jobs are exposed to technological advances in different eras.

We link patent texts to two data sources to identify the two distinct dimensions of innovation as they relate to occupations. To identify innovations that *complement* the output of occupations and industries, we use the tens of thousands of occupational and industry micro-titles supplied by each decade's CAI as a textual corpus characterizing each macro occupation and industry. We refer to these output-complementary patents as *augmentation* innovations. To identify innovations that *substitute* for the inputs of occupations, we use the *Dictionary of Occupational Titles, Revised Fourth Edition*, DOT hereafter (U. S. Department of Labor, Employment and Training Administration, 1991), which describes the tasks accomplished by each occupation.[17] We refer to these input-replacing patents as *automation* innovations.

Figure 4 provides a schematic overview of this process, which we outline in detail in the remainder of this section. We stress that this procedure for linking these two data sources (the CAI and DOT) to patent texts is entirely symmetric and places no structure on the semantic content of the source documents. The degree to which these exercises identify different sets of patents with distinct economic content (i.e., augmentation, innovation) is entirely attributable to differences in the text corpora (CAI vs. DOT).

To link patents to the CAI text corpus, we create a numerical representation of the textual content of each patent and the set of CAI titles falling under a Census occupation (and/or industry), and we use these representations to measure textual similarity. A common approach for comparing textual similarity is to represent documents as vectors that count the number of times a given word shows up in the document; textual similarities are then computed by taking the cosine similarity of these vector representations, relying on exact overlap in terms (what is known as the bag of words approach). As discussed in Kogan et al. (2019), the bag of words method for determining document similarity neglects synonyms and is likely to perform poorly in comparing sets of documents that have disparate vocabularies, as is the case when comparing patent texts with lists of CAI titles. We instead follow

[17]To avoid capturing occupational outputs, we only use task descriptions from the DOT, purging any occupational titles from these descriptions.

Kogan et al. (2019) in representing documents as weighted averages of word embeddings, which overcome the synonym-blindness problem.[18] Word embeddings (Mikolov et al., 2013) are vector representations of word meanings, with highly related words having high cosine similarities between their word embeddings. To turn each word into its vector representation we use the pre-estimated set of word embeddings from Pennington et al. (2014).

For each document, we first clean the text by removing common "stop words" with little informative content, retain all nouns and verbs, and lemmatize each word by converting verbs to their present tense and nouns to their singular form. We then extract the word embeddings for each term in the cleaned document and average across them, leaving us with a vector representation of the document's meaning. We use term frequency–inverse document frequency (TF-IDF) scores to weight the averages.[19]

We call the resulting TF-IDF weighted average of word embeddings a "document vector", which we calculate for all CAI industry or occupation titles[20] for each Census year in our sample and for all United States utility patents issued from 1930–2018. We compute the matrix of cosine similarity scores of patent-occupation pairs. To account for the fact that some types of patents have naturally low similarity scores (e.g. those using highly technical terminology such as chemical patents), we normalize these scores by subtracting the median score across occupations (or industries) for a given patent. For each Census year we restrict the comparison to the CAI titles from that year and the set of patents to the decade preceding that particular Census year. We then retain the top 5 percent highest adjusted textual similarity scores across patent $\times$ occupation pairs as matches for patent p and occupation j:

$$I_{p,j} = 1 \text{ if } X_{p,j} \geq \sigma_\tau, \text{ and zero otherwise.}$$

where $X_{p,j}$ is cosine similarity between patent p and occupation j, and σ_τ is the 95th percentile of the similarity score distribution for period τ.A period τ corresponds to a Census

[18] A "document" is either the full text of a particular patent or the set of CAI titles falling under a Census occupation (or industry) for a given Census year.

[19] TF-IDF weighting of terms is a common approach in textual analysis. The TF-IDF weight of term i in document k is given by $w_{i,k} \equiv TF_{i,k} \times IDF_k$ where $TF_{i,k}$ is the number of times term i occurs in document k divided by the total number of terms in document k, and $IDF_k = \log\left(\frac{\text{N documents in sample}}{\text{N documents that include term } k}\right)$. Thus TF-IDF weighting down-weights terms that occur frequently across documents and up-weights terms that occur frequently within a document. We compute TF-IDF weights separately for patent documents and CAI titles.

[20] Our results are robust to excluding any new industry and occupation titles from the CAI documents prior to patent matching.

year and also the set of patent issue years we consider for that Census year. Typically this will be the previous 10 issue years (so for the 1940 Census τ will consist of patents issued 1930-1939).[21] We find that the method does quite well at identifying matches.[22] Lastly, we take the citation-weighted sum over patents patents issued in period τ to obtain patent counts by occupation over time:

$$\text{Npatents}_{j\tau} = \sum_{p \in \tau} \omega_p \times I_{p,j} \quad \text{with} \quad \omega_p \equiv \tfrac{\text{Ncites}_p}{\text{AvgNcites}_{k(p)}}$$

where $k(p)$ denotes the issue year cohort of patent p. Thus citation weights ω_p are defined as the number of citations received by each patent divided by the average number of citations for patents issued in the same year.[23] We refer to these patents as 'occupation-linked'. We apply the same procedure to construct patent counts by industry i. We further construct occupational exposure to patents linked to industries by weighting industry by occupational employment shares:

$$\text{Npatents}_{j\tau} = \frac{\sum_i E_{ij\tau} \times \text{Npatents}_{i\tau}}{\sum_i E_{ij\tau}} \tag{1}$$

where $E_{ij\tau}$ denotes employment in occupation j and industry i in the Census year for period τ. We refer to these linkages as 'industry-linked' patents. Lastly, when studying occupation-by-industry cell-level outcomes such as employment and wagebill, we use patents linked to both occupation and industry cells: we refer to these linkages as 'industry-occupation-linked' patents. An occupation-by-industry cell, (i,j), is linked to a patent p if the average of the adjusted occupation-patent similarity score (X_{ip}) and industry-patent similarity score (X_{jp}) is among the top 5 percent highest adjusted textual similarity scores across all patent $\times$ occupation $\times$ industry cells.[24]

Summarizing, we have two measures of occupations' exposure to augmentation: the number of direct occupation-patent textual matches ('occupation-linked' patents), and oc-

[21] When we do analysis for time-consistent occupation definitions in the post-1980 period we skip Census year 2010; therefore in this case τ corresponds to patent issue years 2000-2017 for the 2018 Census year.

[22] Appendix Tables A4 and A3 provide some examples of linked patents for Census occupations and industries.

[23] Our results are qualitatively identical when we do not use citations weights and simply sum linked patents.

[24] Due to the large number of (industry $\times$ occupation $\times$ patent) cells, we use a 5% sample of patents to approximate the 95th percentile threshold of the average industry-patent and occupation-patent scores. For example, the 2008-2018 period has over 150 billion (industry $\times$ occupation $\times$ patent) cells. The 95th percentile threshold is calculated using only industry-occupation pairs with non-zero employment counts.

cupational exposure to industry-patent textual matches.[25] Additionally, we combine both sets of textual linkages in 'industry-occupation-linked' patents. In our baseline models, we link patents to occupations and industries over the same time period τ that we measure new work emergence. For example, we link patents awarded between 1930 and 1940 to micro industry and occupation titles in 1940, since we measure new titles in 1940 by comparing Census Alphabetical Indices between 1930 and 1940. As a robustness check, we have also considered specifications where patents are lagged by ten years relative to when new work emergence is observed.

We construct an analogous *automation* exposure measure that identifies technologies that may automate existing labor-using job tasks. We construct this measure identically to above, but replacing CAI micro titles with occupational task descriptions from the DOT. Our procedure for measuring automation innovations follows closely on (Kogan et al., 2019; Webb, 2020), which use the textual similarity between occupational task content and patent texts to measure the ability of new technologies to perform the same work done by workers in particular occupations. Although these augmentation and automation measures are constructed using fully parallel procedures, we demonstrate below that they differ substantially in their predictive relationships with new work emergence and occupational demand shifts.

3 The Shifting Locus of New Work, 1940 – 2018

The occupational distribution of new work has changed markedly over the past eight decades, and these developments mirror—and we believe contribute to—the changing shape of employment growth and skill demands during this period.

As Figure 2 reveals, new work creation was in net weighted heavily towards high-skill occupations (right-hand side of the graph) over the eight decades of our sample. Not evident from this figure, however, is that the rightward skew of new work emergence is a comparatively recent development. This is illustrated in Figure 5, which plots the occupational

[25] As an alternative we have followed Kogan et al. (2017) by assigning patents to industries directly, using patent awarded to firms in Center for Research in Security Prices records linked to Compustat data on the industry of these firms: this eliminates the need for mapping textual similarity. While our results also hold when using these linkages, this approach is not ideally suited to our aims because 1) it doesn't allow us to distinguish between augmentation and automation patents; 2) it restricts patents to those awarded to publicly traded firms, and 3) it implicitly assumes patents are only relevant for the industry where the patent originates, rather than for all industries in which the innovation may be applicable.

distribution of employment in new work separately for 1940–1980 and 1980–2018.[26] Between 1940 and 1980, new worked emerged most strongly in middle-paid production and clerical occupations. After 1980, new work arose disproportionately in high-paid technical, professional, and managerial jobs and, to a lesser extent, in low-paid personal and health services occupations. This observation proves crucial because it links the changing locus new work creation to evolution of skill (i.e., education) demands across the two halves of our sample—as well as the shifting locus of innovation across decades, as suggested by Figure 3.

To assess the educational skew of new work, Figure 6 plots the flow of new work and the stock of preexisting work separately for workers with a high school degree or lower education ('non-college educated') and for workers with at least some college education ('college educated').[27] Among jobs held by non-college workers, there is a sharp divergence in the flow of new work between the two halves of the sample (Panel A). Between 1940 and 1980, new non-college work emerged where non-college workers were already concentrated, that is, in the middle of the occupational pay distribution. Four middle-paid occupations—construction, transportation, production, and clerical and administrative positions—in particular accounted for more than 50 percent of *both* the stock of preexisting non-college work *and* the flow of new work in these four decades.

Over the next four decades, the flow of non-college work diverged from the stock of preexisting work. New non-college work emerged in lower-paid health and personal services, and to a lesser degree in higher-paid sales and professional occupations, even as the stock of non-college work remained relatively concentrated in middle-skill occupations. Two examples illustrate: personal services accounted for 17 percent of the flow of new non-college work between 1980 and 2018 versus only 8 percent of the stock of non-college work; similarly, low-paid health services accounted for 8 percent of the flow of new non-college work versus only 2 percent of the stock of preexisting work. Since over the long run, the flow of new work becomes the stock of existing work, this pattern implies that new work is drawing non-college workers towards lower-paid jobs.

[26] The bars sum to one in each time period.

[27] We calculate the occupational employment of each education group across all Census macro-occupations (approximately 300) in each decade, and then allocate employment within each macro-occupation into new and preexisting work in proportion to the share of titles in that occupation that are newly emergent in that decade. Aggregating these counts into the 12 broad Census occupation categories that can be consistently defined over the entire 1940–2018 period provides the distributions plotted in Figure 6, with each set of bars representing the average occupational distribution by education group during the corresponding time interval.

Panel B of Figure 6 presents complementary evidence for college-educated workers. Between 1940 and 1980, 70 percent of both the stock and the flow of college work was found in three broad occupations: high-skill professional occupations, high-skill managerial occupations, and middle-skill clerical occupations. From 1980 forward, the flow of new work employing college-educated workers skewed even further into the high-skill domain: rising sharply in professional and technical occupations and declining in clerical and administrative work. This pattern augurs a further concentration of college-educated workers in high-skill occupations.

These high level patterns offer two insights: First, the emergence of new work over the last four decades has lead the overall polarization of occupational structure documented by Autor (2019); Autor et al. (2006); Goos et al. (2014); concretely, the flow of new non-college work has shifted more rapidly than has the stock of existing work towards traditionally low-paid low-paid personal service and health-aide occupations, and away from traditionally middle-paid production, clerical, and administrative support occupations. Second, and by implication, employment polarization does not merely reflect an erosion of employment in existing middle-skill work but also a change in the locus of new work creation. As the emergence of new types of non-college work has slowed in middle-paid occupations and accelerated in low-paid occupations, the allocation of non-college workers across occupations has tracked the shifting locus of new work emergence.[28]

4 Theoretical Framework

To develop intuition about the evolution of new work and guide the empirical analysis, we offer a model, building on Acemoglu and Restrepo (2018), Acemoglu and Autor (2011) and Autor et al. (2003), that considers how three forces shape the endogenous creation of new job tasks and the elimination of old tasks: augmentation, which generates new labor-using job tasks; automation, which eliminates labor-using tasks; and shifts in consumer demand which affect task automation and new task creation by changing innovation incentives. This model provides predictions about when and where new tasks emerge, how the flows of augmentation and automation innovations correlate across occupations, and how these two distinct faces of innovation affect (and are jointly determined with) occupational labor demand.

[28]This pattern is not automatic. Non-college workers could reallocate from middle- to low-paid occupations without any change in the locus of new work creation.

4.1 Environment

We begin with two sectors, producing skill-intensive and skill-non-intensive goods or services, Y_S and Y_U. The subscripts denote the respective sectors. A representative household consumes goods Y_U and Y_S according to:

$$U(Y_U, Y_S) = Y_U^{\beta_U} Y_S^{\beta_S}, \tag{2}$$

where $\beta_U + \beta_S = 1$; P_j is the price of good j with $j = U, S$; P the ideal price index; Y is total utility; and $P_U Y_U + P_S Y_S = PY$. Let Y be the numéraire so that $P \equiv 1$.[29] We will later allow β to be change to reflect demographic forces that shift preferences for consumption between skill-intensive and skill non-intensive services. We simplify the structure of consumption by assuming that there is no leisure and hence labor supply is inelastic.

Each sector produces a unique final output by combining a unit measure of tasks $i \in \{N_j - 1, N_j\}$:

$$Y_j = \left[\int_{N_j-1}^{N_j} y_j(i)^{\frac{\sigma-1}{\sigma}} di\right]^{\frac{\sigma}{\sigma-1}} \tag{4}$$

where $y_j(i)$ is the output of task i in sector j; σ is the elasticity of substitution between tasks (assumed identical across sectors $j \in \{U, S\}$); and $A_j > 0$ is a constant.

Each task is produced by combining labor composite of high- and low-skill types, $n_j(i)$, or capital, $k_j(i)$ with a task-specific intermediate $q_j(i)$. The production function for task i is given by:

$$y_j(i) = \begin{cases} B_j q_j(i)^\eta k_j(i)^{1-\eta} \text{ if } i \in [N_j - 1, I_j] \\ B_j q_j(i)^\eta [\gamma_j(i) n_j(i)]^{1-\eta} \text{ if } i \in (I_j, N_j] \end{cases} \tag{5}$$

where $B_j \equiv \psi_j^\eta [1-\eta]^{\eta-1} \eta^{-\eta}$ for notational convenience; the parameter $\eta \in (0, 1)$ is the share of output paid to intermediates; $\gamma_j(i)$ is the productivity of the labor composite $n_j(i)$ (relative to capital); and I_j and N_j are the equilibrium thresholds for automation and new task creation, respectively, meaning tasks from $N - 1$ to I_j are produced by machines and those from I_j to N_j are produced by labor. We make the following assumption:

[29] Given that consumption is Cobb-Douglas, P is given by:

$$P(P_U, P_S) = \left[\frac{P_U}{\beta_U}\right]^{\beta_U} \left[\frac{P_S}{\beta_S}\right]^{\beta_S} = 1 \tag{3}$$

Assumption 1 $\gamma_j(i)$ *is strictly increasing.*

Assumption 1 implies that in each sector, labor has strict comparative advantage in tasks with a higher index. This assumption guarantees that, in equilibrium, tasks with lower indices will be automated in each sector, while those with higher indices will be produced with labor. Due to this strict comparative advantage structure, there is a unique threshold $\tilde{I}_j$ in each sector such that $\frac{W_j}{R_j} = \gamma_j\left(\tilde{I}_j\right)$, where $W_j \equiv W_L^{\alpha_j} W_H^{1-\alpha_j}$, W_L and W_H equal the economy-wide wage for L and H labor, and R_j is the rental rate for sector-specific capital.

Task-specific intermediate $q_j(i)$ embodies the technology used either for automation or for the creation of new labor intensive tasks. We start by assuming that these intermediates are supplied competitively and that they can be produced using ψ_j units of the sector-specific good. Hence, they are also priced at ψ_j in units of sectoral output. (In section 4.4 we additionally model endogenous innovation responses.) The measures of high-skill and low-skill labor are given by $H > 0$ and $L > 0$, respectively. The labor composite $n_j(i)$ in each sector is a Cobb-Douglas combination of H and L labor:

$$n_j(i) = l_j(i)^{\alpha_j} h_j(i)^{1-\alpha_j}. \tag{6}$$

Both types of labor are used in each sector, but H labor is used more intensively in the more skill-intensive S sector, and L labor is used more intensively in the skill-non-intensive U sector ($0 < \alpha_S < \alpha_U < 1$). Let L_U, L_S, H_U, and H_S be the equilibrium labor allocations to each sector. Then, $L_U + L_S = L$ and $H_U + H_S = S$. We define here a wage index reflecting the price of the sectoral labor composite, $W_j \equiv W_L^{\alpha_j} W_H^{1-\alpha_j}$, where W_L and W_H equal the economy-wide wage for L and H labor, respectively. Finally, capital is sector-specific, with sectoral capital stocks K_U and K_S taken as given, and R_j is the capital rental rate for sector-specific capital.

4.2 Equilibrium

Before characterizing the equilibrium in our model, we simplify with two assumptions.

Assumption 2 *We have* $K_j < \bar{K}_j$*, where* $\bar{K}_j$ *is such that* $R_j = \frac{W_j}{\gamma(N_j)}$ *for* $j \in \{u, s\}$.

This ensures that the capital rental rate is sufficiently high in each sector that new tasks will be adopted immediately and will increase aggregate output. If Assumption 2 were not satisfied, new tasks would be more expensive to produce than the tasks that they potentially displace, i.e., the lowest index tasks, so that new tasks would either reduce productivity or would simply not be adopted.

The next assumption simplifies the determination of the automation threshold, I_j. Because labor has a strict comparative advantage in tasks with a higher index, i.e. i, there is a unique threshold $\tilde{I}_j$ in each sector such that

$$\frac{W_j}{R_j} = \gamma_j(\tilde{I}_j) \tag{7}$$

For all tasks $i \leq \tilde{I}_j$, we have that $R_j \leq W_j/\gamma_j(i)$, so these tasks are potentially more cheaply produced with capital. However, if $I_j < \tilde{I}_j$, then the state of automation acts as a constraint on which tasks are accomplished by capital. In particular, the threshold task that will be performed by capital is $I_j^* = \min\{\tilde{I}_j, I_j\}$. We simplify the set of cases considered by invoking the following assumption:

Assumption 3 *We have that $I_j^* = I_j \leq \tilde{I}_j$, so that the threshold task in each sector is constrained by the state of automation.*

Assumption 3 implies that when a new automation technology is introduced, it is always adopted. With this assumption and the fact that tasks are competitively supplied, the price of task i, $p(i)$, is given by:

$$p_j(i) = \begin{cases} R_j^{1-\eta} \text{ if } i \in [N_j - 1, I_j] \\ [W_j/\gamma_j(i)]^{1-\eta} \text{ if } i \in (I_j, N_j] \end{cases} \tag{8}$$

Combining equations (2) and (4), the demand for sectoral task output $y_j(i)$ is:

$$y_j(i) = [P_j/p_j(i)]^{\sigma} Y_j = \beta_j Y P_j^{\sigma-1} p_j(i)^{-\sigma} \tag{9}$$

Together with the fact that the supply of $y_j(i)$ is a Cobb-Douglas aggregate of labor, capital, and intermediates, we can obtain the sectoral demands for capital and labor for each task i, respectively:

$$k_j(i) = \begin{cases} [1-\eta]\beta_j Y P_j^{\sigma-1} R_j^{-\hat{\sigma}} \text{ if } i \in [N_j - 1, I_j] \\ 0 \text{ if } i \in (I_j, N_j] \end{cases} \tag{10}$$

and

$$l_j(i) = \begin{cases} 0 \text{ if } i \in [N_j - 1, I_j] \\ [1-\eta]\beta_j Y P_j^{\sigma-1} \frac{1}{\gamma_j(i)} \left[\frac{W_j}{\gamma_j(i)}\right]^{-\hat{\sigma}} \text{ if } i \in (I_j, N_j] \end{cases} \tag{11}$$

We can define a static equilibrium in a similar way to Acemoglu and Restrepo (2018): Given a range of tasks $[N_j - 1, N_j]$, automation technology $I_j \in (N_j - 1, N_j]$, and a capital stock K_j for each sector j, a static equilibrium is summarized by a set of factor prices W_L, W_H, and R_j; threshold tasks $\tilde{I}$ and I^*; employment levels, L_j and H_j; and aggregate output, Y_j, for each sector j, such that

- $\tilde{I}_j$ is determined by equation (7) and $I_j^* = \min\{I_j, \tilde{I}_j\}$, which is equal to I_j by Assumption 3;
- The capital and labor markets clear in each sector, so that

$$\int_{N_j-1}^{N_j} [1-\eta]\beta_j Y P_j^{\sigma-1} R_j^{-\hat{\sigma}} di = K_j \tag{12}$$

$$\int_{N_j-1}^{N_j} [1-\eta]\beta_j Y P_j^{\sigma-1} \frac{1}{\gamma_j(i)} \left[\frac{W_j}{\gamma_j(i)}\right]^{-\hat{\sigma}} di = L_j \tag{13}$$

  where $\sum_j L_j = L$;
- Factor prices satisfy the ideal price index condition:

$$P_j^{1-\sigma} = [I_j - N_j + 1] R_j^{1-\hat{\sigma}} + W_j^{1-\hat{\sigma}} \int_{I_j}^{N_j} \gamma_j(i)^{\hat{\sigma}-1} di. \tag{14}$$

Proposition 1 *In the static equilibrium defined above, aggregate output of sector j is given by:*

$$[1-\eta]Y_j = P_j^{\frac{\eta}{1-\eta}} \left[[I_j - N_j + 1]^{\frac{1}{\hat{\sigma}}} K_j^{\frac{\hat{\sigma}-1}{\hat{\sigma}}} + \left[\int_{I_j}^{N_j} \gamma_j(i)^{\hat{\sigma}-1} di\right]^{\frac{1}{\hat{\sigma}}} L_j^{\frac{\hat{\sigma}-1}{\hat{\sigma}}} \right]^{\frac{\hat{\sigma}}{\hat{\sigma}-1}} \tag{15}$$

Proof See Appendix F.

This model generalizes the single-sector setting in Acemoglu and Restrepo (2018) to two sectors with different skill-intensities. When we consider demand expansions and contractions below, the interaction of these sectors will be useful.

4.3 Innovation and Employment

Having laid out the general model above, we can now consider the consequences of changes in the task structure of labor demand in each sector, specifically, the effects of task automation and task augmentation. Automation occurs when previously labor-using tasks are taken over by capital, corresponding to a rise in the sectoral automation threshold, I_j. Augmentation refers to the introduction of new labor-using tasks in a sector, corresponding to a rise in N_j. In a single-sector model, the effect of augmentation and automation on labor demand depend solely on substitution and scale effects in that sector. In our multi-sector setting with labor mobility and heterogeneous skills, the aggregation of labor and inter-sectoral labor flows are important for the consequences of automation and augmentation on labor demand. Augmentation and automation in either sector will affect labor demand in both sectors and cause labor to reallocate across sectors.

Proposition 2 (Employment effects of automation and augmentation) *Automation in sector U (a rise in I_U) increases the range of sector U tasks produced by capital, which decreases employment of both high-skill and low-skill workers in that sector. These workers move to sector S. Augmentation in sector U (a rise in N_U) has the converse effect: by introducing new labor-using tasks in sector U, it increases employment of both high-skill and low-skill workers in that sector, drawing away these workers from sector S. That is,*

$$\begin{array}{ll} \frac{\partial L_U}{\partial I_U}, \frac{\partial H_U}{\partial I_U} < 0, & \frac{\partial L_S}{\partial I_U}, \frac{\partial H_S}{\partial I_U} > 0 \\ \frac{\partial L_U}{\partial N_U}, \frac{\partial H_U}{\partial N_U} > 0, & \frac{\partial L_S}{\partial N_U}, \frac{\partial H_S}{\partial N_U} < 0. \end{array}$$

These derivatives have the opposite sign when augmentation or automation occurs in sector S.

Proof See Appendix F.

This proposition, a main result of the conceptual framework, reveals the direction of labor flows in response to automation and augmentation. All else equal, automation in a sector leads to the contraction of that sector by reducing employment of both types of workers, whereas augmentation in a sector attracts workers of both types.

Three mechanisms jointly underlie the co-movement of low- and high-skill workers across sectors in response to automation or augmentation. First, tasks are gross substitutes in each sector ($\sigma > 1$), so automation in a given sector implies a fall in that sector's labor share (and conversely for augmentation). Second, demand for high- and low-skill labor in each sector is Cobb-Douglas, so the wagebill paid to each skill group by a sector is proportional to that sector's labor share. Finally, the share of aggregate expenditure devoted to each sector is fixed by the utility function (equation 2). Hence, automation in a sector spurs a decline in the sector's labor share, yielding an inward shift in both high- and low-skill sectoral labor demand relative to the other sector.

This observation is a key input into our empirical work, implying that a sector's employment rises with sector-specific augmentation and falls with sector-specific automation. We test this implication directly in Section 7, where we equate occupations in the empirical analysis with sectors in the model.

Naturally, changes in sectoral labor demands alter the economy-wide skill premium, W_H/W_L, as explained in the next corollary.

Corollary 1 (Sectoral innovations and the aggregate skill premium) *Automation in the U sector raises the skill premium, W_H/W_L, by reducing labor demand in the low-skill intensive sector. Augmentation in the U sector lowers the skill premium by increasing labor demand in the low-skill intensive sector. Conversely, automation in the S sector lowers the skill premium while augmentation in the S sector raises the skill premium. Formally,*

$$\frac{\partial(W_H/W_L)}{\partial N_U}, \frac{\partial(W_H/W_L)}{\partial I_S} < 0, \quad \frac{\partial(W_H/W_L)}{\partial I_U}, \frac{\partial(W_H/W_L)}{\partial N_S} > 0.$$

This corollary spells out general equilibrium implications of innovations that reallocate the distribution of tasks between labor and capital in either sector. Our empirical analysis does not focus on these general equilibrium empirical implications, and the next corollary explains why.

Corollary 2 (Changes in sectoral wagebills by skill group) *Due to the law of one price for skill, the effect of innovation on the log sectoral wagebill of a skill group relative to its wagebill in the non-innovating sector is identical to its effect on the log relative sectoral employment of that skill group. Formally:*

$$\begin{aligned} \frac{\partial \ln(W_L L_U/W_L L_S)}{\partial I_U} &= \frac{\partial \ln(L_U/L_S)}{\partial I_U}, \quad & \frac{\partial \ln(W_L L_U/W_L L_S)}{\partial N_U} &= \frac{\partial \ln(L_U/L_S)}{\partial N_U} \\ \frac{\partial \ln(W_H H_U/W_H H_S)}{\partial I_U} &= \frac{\partial \ln(H_U/H_S)}{\partial I_U}, \quad & \frac{\partial \ln(W_H H_U/W_H H_S)}{\partial N_U} &= \frac{\partial \ln(H_U/H_S)}{\partial N_U}, \end{aligned}$$

and similarly for innovation in the S sector.

This corollary, which echoes Proposition 3 in Hsieh et al. (2019), follows from the mobility of labor across sectors and is not specific to the CES structure of task demand or the Cobb-

Douglas structure of labor demand in each sector. In combination with Proposition 2, this corollary provides a testable implication, which is that the impact of sectoral innovations—which we measure using augmentation and automation patents—on the sectoral wagebill by skill group will mirror those for sectoral employment. We test this implication in Section 7.

4.4 Shifts in Consumer Demand and Innovation

To understand the interaction between shifts in consumer demand and innovation, we work with a simple, one-period framework, which utilizes the general results above but endogenizes the supply of intermediates which embody the task-specific technology. At the start of the period, the parameters determine the equilibrium variables: factor prices and output. A continuum of firms, given such information, hires entrepreneurs of exogenous supply E, where E is some large number. These entrepreneurs can be hired to work in four sector-innovation cells: automation in sector U, new task creation in sector U, automation in sector S, and new task creation in sector S. We denote the number of entrepreneurs in each sector-innovation cell E_I^U, E_N^U, E_I^S, and E_N^S, respectively.

Upon being hired, these entrepreneurs generate new intermediates which embody augmentation and automation technologies according to

$$\Delta I^j = E_I^j \tag{16}$$

$$\Delta N^j = E_N^j \tag{17}$$

ΔI^j and ΔN^j are realized immediately.

Entrepreneurs have utility given by

$$U_{z,m}^j = \max_{\{m,j\}} \{w_m^j + \nu \epsilon_{z,m}^j\} \tag{18}$$

where $U_{z,m}^j$ is the (period) utility of entrepreneur z hired to work on innovation m in sector j. The idiosyncratic preference terms $\epsilon_{z,m}^j$ are independent Type-I Extreme Value draws with zero mean, and the parameter ν scales the variance of these idiosyncratic terms. Entrepreneurs choose the sector and innovation activity that delivers the highest utility.

Under the distributional assumptions above, the share of entrepreneurial labor supplied to each sector-innovation cell has a closed-form analytical expression. Denote by π_m^j the fraction of entrepreneurs that move to sector j to work on innovation m. Then,

$$\pi_m^j = \frac{\exp(w_m^j)^{1/\nu}}{\sum_m \sum_j \exp(w_m^j)^{1/\nu}} \tag{19}$$

Thus, $1/\nu$ can be interpreted as a labor supply elasticity, as in Caliendo et al. (2019). Applying the law of large numbers, the measure of entrepreneurs hired in sector j working on innovation m is

$$E_m^j = \pi_m^j E \tag{20}$$

Competition among prospective technology monopolists to hire entrepreneurs implies wages as follows:

$$w_m^j = V_m^j \tag{21}$$

where V_m^j is the value of innovation m in sector j.

Demand for intermediate $q_j(i)$ is given by:

$$q_j(i) = \begin{cases} \psi_j^{-1}\eta Y_j P_j^\sigma R_j^{(1-\eta)(1-\sigma)} \text{ if } i \in [N_j - 1, I_j] \\ \psi_j^{-1}\eta Y_j P_j^\sigma \left(\frac{W_j}{\gamma_j(i)}\right)^{(1-\eta)(1-\sigma)} \text{ if } i \in (I_j, N_j] \end{cases} \tag{22}$$

where we can also substitute $Y_j P_j^\sigma = \beta_j Y P_j^{\sigma-1}$. Gross profit from automating task I are

$$\pi(I) = \begin{cases} (1-\mu)\psi_j q_j(I) = (1-\mu)\eta Y_j P_j^\sigma R_j^{(1-\eta)(1-\sigma)} & \text{if } I \text{ is produced with capital} \\ (1-\mu)\psi_j q_j(I) = (1-\mu)\eta Y_j P_j^\sigma \left(\frac{W_j}{\gamma_j(I)}\right)^{(1-\eta)(1-\sigma)} & \text{if } I \text{ is produced with labor} \end{cases} \tag{23}$$

Note that $(1-\eta)(1-\sigma) \equiv 1-\hat{\sigma}$. Hence the sectoral value of automating task I and N are given by, respectively:

$$V_j^I = (1-\mu)\eta Y_j P_j^\sigma \left[R_j^{1-\hat{\sigma}} - \left(\frac{W_j}{\gamma_j(I)}\right)^{1-\hat{\sigma}} \right] \tag{24}$$

$$V_j^N = (1-\mu)\eta Y_j P_j^\sigma \left[\left(\frac{W_j}{\gamma_j(N)}\right)^{1-\hat{\sigma}} - R_j^{1-\hat{\sigma}} \right] \tag{25}$$

Having determined the sectoral value of automating task I and N, we study how these incentives for automation and new task creation change in sector j in response to a demand expansion, i.e., an increase in β_j.

Lemma 1 *In equilibrium, we have that $V_j^N = V_j^I$.*

Lemma 1 implies that entrepreneurs are initially indifferent between creating new automation or augmentation intermediates: if this were untrue, the allocation of entrepreneurs across innovation margins is not in equilibrium. Note that in this equilibrium there are still positive productivity gains from additional task automation and from additional new task creation, by Assumptions 2 and 3.

As these incentives given by V_j^I and V_j^N determine wages, the employment share and changes in I and N for each sector naturally follow by consulting (19) and the innovation production functions (16) and (17).

Proposition 3 *A demand shift towards a given sector unambiguously increases new task creation relative to automation in that sector, while decreasing new task creation relative to automation in the other sector.*

$$\frac{\partial \Delta N_j}{\partial \beta_j} > \frac{\partial \Delta I_j}{\partial \beta_j}$$
$$\frac{\partial \Delta N_{\bar{j}}}{\partial \beta_j} < \frac{\partial \Delta I_{\bar{j}}}{\partial \beta_j}$$

Proof. See Appendix F.

The proposition indicates a positive relationship between demand shifts and new task creation. When there is a positive demand shift in a given sector j, the incentives for new task creation in that sector increase as a result of movement on two margins: on the demand side, both output and price increases, and on the factor side, the price of capital increases more than that of effective labor as capital supply is inelastic, increasing the price differential between the two factors. This increased price differential raises the potential returns to new task creation, which assigns tasks from capital to labor.[30] In Section 6 we test an implications of this Proposition, specifically, that outward demand shifts accelerate new task emergence whereas inward demand shifts decelerate it.

[30] This result relies on the fact that tasks are gross substitutes in each sector, $\sigma > 1$, so that a change in the sectoral relative price of capital versus labor increases the profitability of innovations that expand usage of the factor whose relative price has fallen.

5 Augmentation exposure and new task creation

Following the logic of the model, this section empirically characterizes the forces that explain where and when new job tasks emerge by relating the emergence of new occupational tasks to the exposure of occupations to: (a) augmentation; (b) automation; and (c) demand shifts. Our focus here is on forces that affect the creation of *new job tasks*, meaning the emergence of new titles. We take up the net effects of new work creation on occupational employment in section 7.

5.1 Where do new tasks emerge?

In our conceptual framework, economic forces that complement occupational outputs lead to the emergence of new work tasks. One of those forces is augmentation. The first hypothesis that we test is that new job tasks emerge differentially in occupations that are more exposed to augmentation innovations—meaning those that may complement occupational outputs. We estimate models of the following form:

$$\text{Newtitles}_{jt} = \beta_1 \text{AugX}_{jt} + \beta_2 \frac{E_{jt}}{\sum_j E_{jt}} + D_t \left(+D_J + D_{Jt}\right) + \varepsilon_{jt}, \tag{26}$$

where j indexes Census occupations, and t indexes decades (1930–1940, 1940–1950, .., 2000–2010, 2010–2018). The dependent variable is a measure of the flow of new work titles occurring in a Census occupation in a decade, and the independent variable of interest is AugX_{jt}, measuring occupational exposure to augmentation as revealed by textual links between utility patents and the *Census Alphabetical Index of Industries and Occupations*. Both variables in year t are measured as cumulative flows over the preceding decade: new work observed in 1940 has emerged over 1930–1940; and similarly, augmentation exposure in 1940 is constructed from patents awarded over 1930–1940. Year fixed effects absorb year-by-year variation in the total number of new titles. We control for the employment (E) share of occupation j to remove any mechanical association between new title counts and relative occupational employment size. In some specifications, we further add fixed effects for the twelve consistently-defined broad occupational groups, indexed by J, and their interaction with year fixed effects.

The dependent variable, the count of new titles in a macro-occupation in a decade, contains many zeros. We apply two functional forms that handle these zeros slightly differently. In Panel A of Table 2, we use as our dependent variable the inverse hyperbolic sine (IHS) of new titles and, similarly, use the IHS of the count of augmentation patents as our explanatory variable. In Panel B, we instead use year-specific percentiles of both the outcome and explanatory variables. As we document, our core findings are consistently robust to these variations.[31]

It is infeasible to construct a fully balanced panel of detailed occupations over the eight decades between 1940 – 2018 without sacrificing substantial resolution. We accordingly employ an unbalanced approach for models that pool these eight decades. (When estimating models for the second half of our sample, 1980 through 2018, we use a balanced panel approach based on (Dorn, 2009).) In each decade, we use the full set of occupational categories available in the corresponding Census, while employing the twelve broad, consistent occupational categories seen in earlier figures. This enables us to compare new title emergence across all occupations over eight decades, or across all occupations within 12 broad categories across decades, or across occupations within 12 broad categories in each decade.[32] Because our independent and dependent variables capture within-occupation flows—new occupational titles added during a decade, new occupation-related patents issued during that decade—each decade of data is implicitly a 10-year occupation-level panel.

Based on our hypotheses and theoretical framework, we expect $\beta_1 > 0$: more augmentation-exposed occupations will add more new titles. The first three columns of Table 2 report estimates of equation (26) using industry-linked patents—meaning patents linked to occupations according to their distribution across more vs. less augmentation-exposed industries. The first column estimate of 0.159 (se = 0.034) implies that each 10 percent increment to augmentation exposure predicts an additional 1.6 percent higher rate of new title emergence over the course of a decade. In column 2, we add 12 broad occupation dummies, thus comparing rates of new title emergence across detailed occupations within broad occupation categories. The point estimate is 0.115 in this specification, and remains highly precise. The third column adds interactions between the 12 occupation dummies and decade effects, thus further limiting the comparison to new title emergence rates across detailed occupations within broad occupational categories within each decade. These additional controls absorb

[31] We have also experimented with using a binary measure of new titles (zero/non-zero) and with using the logarithm of new titles conditional on positive. These approaches qualitatively verify our core results.

[32] The specificity of Census 3-digit occupational categories (as well as industry categories) generally rises from decade to decade, with approximately 250 3-digit occupations in 1940 and roughly 500 in 2018. Occupation and industry categories also merge, split, and recombine, so harmonization comes at a high cost in foregone resolution.

significant additional variation in the outcome variable, as seen from R-squared value (0.754 in column 3 versus 0.674 in column 2). The point estimate remains highly similar to the prior column (0.096, SE = 0.018), nevertheless.

In the second set of three columns in Table 2, we measure augmentation exposure by using occupation-linked patents rather than patents associated with occupations via their industry weights. We obtain point estimates that are nearly identical to those for the industry-linked patents with slightly higher precision. The addition of major occupation dummies (column 2) and occupation-by-decade dummies (column 3) has essentially no effect on the point estimates.

To check the sensitivity of these estimates to the use of the IHS transformation of the dependent and independent variables, Panel B of Table 2 re-estimates these models replacing the IHS with year-specific percentiles of new titles (dependent variable) and patent exposure (independent variable).[33] These models yield quantitatively large and statistically precise point estimates. For example, the column 3 and column 6 point estimates imply that a 10 percentile increase in augmentation exposure predicts an additional 0.95 to 3.60 percentile increase in new title emergence per decade. Comparing an occupation at the 75th versus 25th percentile of augmentation exposure in a decade, we would expect the more exposed occupation to fall 5 to 18 percentiles higher in the new title emergence distribution in that decade.

5.2 Contrasting automation with augmentation

A second implication of our conceptual framework is that technologies that automate existing occupational tasks should *not* spur the emergence of new occupational tasks. We test this implication by incorporating our measure of occupational exposure to automation innovations alongside the augmentation exposure measure above. Since both measures are built from the same corpus of patents, we view this as a (joint) test of the capacity of our NLP procedure to distinguish automation from augmentation patents, and of the model's implication that automation technologies do not tend to generate new occupational labor-using tasks.

Figure 7 previews the substantive content of the automation exposure measure by plotting the bivariate relationship between percentiles of automation exposure and percentiles of aug-

[33] Year-specific percentiles account for the fact that the number of patents awarded annually rises across decades.

mentation exposure at the level of 303 three-digit ('macro') occupations.[34] We focus on the most recent four decades of our sample, 1980–2018, where we can form a consistently defined panel of of three-digit Census occupations and industries.[35] Occupations exposed to more augmentation are also more exposed to more automation. The employment-weighted cross-occupation correlation between augmentation and automation exposure is 0.62. This positive correlation is logical. Many technologies contain both automation and non-automation components.[36] It is also consistent with our theoretical framework, which highlights that incentives to automate and innovate are in part responding to the same profit opportunities.

Alongside the strong positive correlation, the off-diagonal occupations are instructive. Typesetters and compositors, Clinical laboratory technologists and technicians, Cabinetmakers and bench carpenters, and Machinists are four occupations that have a high rate of automation relative to augmentation. Our conceptual framework predicts that employment in these occupations would tend to erode. Conversely, Mechanical engineers, Operations and systems researchers and analysts, and (to a lesser degree) Supervisors of mechanics and repairers, and Business and promotion agents are occupations where augmentation has outpaced automation. We would expect these occupations to expand. Finally, the on-diagonal examples illustrate that occupations may have a relatively 'balanced' degree of exposure to both automation and augmentation either because they are highly subject to both forces (e.g., Assemblers of electrical equipment) or because they are relatively insulated from both (e.g., Clergy and religious workers).[37]

To test whether automation exposure does *not* spur the emergence of new job tasks, we report in Table 3 estimates of equation (26) in the 1980–2018 panel, expanded with measures of both augmentation and automation exposure. Models include year main effects and control for occupational employment shares to avoid a mechanical relationship between occupation size and the rate of new title emergence. The first three columns of Panel A

[34] Both augmentation and automation exposure are averaged over 1980–2018 for each occupation.

[35] This harmonization adapts the classifications in Dorn (2009) and Deming (2017), further updated to encompass Census occupations and industries through 2018.

[36] Indeed, 55 percent of the total occupation-linkages of patents to Alphabetical Index occupation titles are also linked based on occupational task content

[37] Appendix Table A9 documents the correlation between these augmentation and automation measures and task measures commonly used to capture the potential for technology to complement or substitute workers (Autor and Dorn, 2013; Autor et al., 2003). Occupations' automation exposure is significantly positively correlated with routine-task content, whereas the converse holds for augmentation exposure. We jointly include augmentation and automation in these models, highlighting each has substantial independent predictive variation for widely-used task measures.

of Table 3 re-estimate our baseline specifications, which predict new title emergence as a function of occupational augmentation exposure. Augmentation exposure is measured using occupation-linked patents for maximum comparability to the automation measure, which is also based on occupation linkages.

Column 1 contains the basic specification, while columns 2 and 3 add respectively broad occupation fixed effects and broad occupation-by-year interactions. We obtain a statistically precise point estimate on AugX of around 0.15 in all three columns (with a standard error of around SE = 0.02), in line with our previous findings over 1940–2018. Column 4 of Table 3 replaces augmentation exposure with automation exposure, measured as patents linked to occupational tasks. When *not* controlling for augmentation exposure (as in this column), automation exposure is strongly positively associated with new title emergence with a point estimate of 0.105 (SE = 0.033). However, when both augmentation exposure and automation exposure are included (columns 5–7), only the augmentation exposure predicts new title emergence. Conditional on augmentation exposure, automation exposure is not associated with new title growth: point estimates are small and negative, as well as statistically insignificant. This pattern is striking since augmentation and automation exposure are close statistical relatives as show in Figure 7. While a skeptical reading of these results is that AutomX contains no economic content beyond what is already present in AugX, we show in Section 7 that both variables have significant independent (and oppositely-signed) predictive power for occupational employment growth.

We demonstrate the robustness of this pattern in Panel B of Table 3. Here, we re-specify the independent and dependent variables as occupational percentiles of new title counts, automation exposure, and augmentation exposure (replacing their IHS counterparts in Panel A). Columns 1–3 find that the augmentation exposure point estimates for 1980–2018 are highly comparable to those for 1940–2018. As above, when not controlling for augmentation exposure (column 4), automation exposure is predictive of new title emergence, with a point estimate of 0.122 (SE = 0.067). Columns 5–7 show that this positive relationship is entirely eliminated when AugX is included. A 10 percentile higher augmentation exposure rank predicts approximately a 4.7 percentile higher new title count rank. The point estimate on automation exposure is negative and statistically insignificant when the control for augmentation exposure is included.

6 Demand Shifts and New Task Creation

A central implication of our conceptual framework is that the emergence of new work responds elastically to market size: positive demand shifts foster the emergence of new labor tasks by raising the value of occupational outputs and spurring augmentation innovations; negative demand shifts hinder the emergence of new labor tasks by lowering the value of occupational outputs, thus deterring augmentation innovations. We construct two demand shift measures to study these relationships. To isolate negative demand shifts, we exploit an exogenous decline in industry-level manufacturing demand resulting from rising Chinese import competition over 1990–2018. To isolate positive demand shifts, we leverage changes in population age structure that indirectly affect employment through patterns of consumption, following DellaVigna and Pollet (2007).[38] In both cases, we measure the differential exposure of occupations to demand shifts as a function of their distribution across more-exposed versus less-exposed industries. Although both inward and outward demand shifts are predicted to affect the arrival rate of new work, we find it useful to apply countervailing tests since new work never flows in reverse and occupational titles are hardly ever removed from the Census Index. We therefore test whether negative demand shocks slow the emergence of new titles rather than whether they spur the elimination of of new or existing titles.[39]

6.1 Demand contractions and new task deceleration

Starting in the early 1990s, import competition from China generated a sizable negative demand shock for many labor-intensive domestic manufacturing industries in the U.S. (Autor et al., 2014; Bernard et al., 2006; Pierce and Schott, 2016).[40] We use these industry-level demand shocks to identify shocks to occupational labor demand. We construct occupation

[38] Our work is also related to Comin et al. (2020); Leonardi (2015); Mazzolari and Ragusa (2013), who analyze the causal effects of age, education, and income on employment via consumption. Appendix D.1 details the construction of our demand shift measures.

[39] Because our two demand measures cover different time intervals at different periodicities, we do not currently pool them in a single estimating equation.

[40] Related work maps trade shocks to labor market outcomes in Brazil, Canada, India, Norway, Germany, Mexico, and other countries (Balsvik et al., 2015; Branstetter et al., 2019; Chiquiar, 2008; Dauth et al., 2014; Devlin et al., 2021; Kovak, 2013; Topalova, 2010).

j's change in exposure to Chinese import competition, ΔCI_{jt}, as follows

$$\Delta \mathrm{CI}_{jt} = \sum_i \frac{E_{ij,t-1}}{E_{j,t-1}} \times \frac{\Delta M_{i,t}^{OC}}{Y_{i,88} + M_{i,88} - X_{i,88}}, \tag{27}$$

where $\Delta M_{i,t}^{OC}$ are changes in industry i's imports from China by a set of developed countries other than the United States over the periods 1991–2000 and 2000–2014. $Y_{i,88} + M_{i,88} - X_{i,88}$ is initial absorption of the U.S. industry, measured as the real value of industry shipments plus industry imports minus industry exports, all measured in the initial year 1988. By using China's industry-level exports to non-U.S. destinations as our predictor of import shocks to United States industries, we are implicitly using a reduced form approach to measuring U.S.-facing import competition shocks stemming from China's rise as an exporter (Acemoglu et al., 2016; Autor et al., 2014; Jaravel and Sager, 2020). We obtain occupational exposure by multiplying changes in predicted U.S. industry import shocks by the occupation's employment share across US industries in the initial year. While this measure is only available for 1990 onward (reflecting changes in import competition over 1991–2000 and 2000–2014), the advantage is that it captures a plausibly exogenous decline in domestic occupational demand stemming from external forces.[41]

China trade exposure varies substantially both within and between production and non-production occupations, as shown in Figure 8, which plots percentiles of occupational exposure to import competition over 1990–2018 for occupations classified into broad occupational groups. Broad occupations are ordered according to their average China trade exposure on the y-axis, while variation across detailed occupations within these 12 broad categories is depicted along the x-axis. In particular, within production occupations, Power plant operators have a trade shock exposure of only around the 40th percentile of the average across all occupations, whereas Textile sewing machine operators are the most exposed. In transportation occupations, Bus drivers are relatively unexposed, as are Insurance adjusters among clerical occupations, and Primary school teachers among professionals. On the other hand, Machine feeders and offbearers in transportation occupations; Shipping and receiving clerks in clerical occupations; and Electrical engineers in professional occupations are more highly exposed than the average occupation within production.

We test whether this demand shift affects the emergence of new occupational titles using

[41] We scale these changes to match 1990–2000 and 2000–2018, the time periods we observe in our new title data.

the following specification

$$\text{IHS}\left(\text{newtitles}_{jt}\right) = \beta_1 \text{ImportX}_{jt} + D_t + Z_{jt} + \varepsilon_{jt}, \tag{28}$$

where ImportX_{jt} are year-specific occupational percentiles of ΔCI_{jt}, and Z_{jt} is a vector of controls. This vector includes occupational exposure to augmentation and overall occupational employment shares, as before, but also occupations' employment shares across all 13 broad industries. This means that β_1 is identified by variation in occupation-level trade exposure stemming from differences in occupations' distributions across manufacturing and non-manufacturing industries. In some specifications, we additionally control for broad occupation fixed effects as well as occupational employment growth to remove any potential mechanical association between employment contractions and declines in new title emergence. With occupation fixed effects included, identification stems from contrasts across detailed occupations within the twelve broad occupational categories depicted in Figure 8—in effect, from variation within rows of Figure 8. Our expectation is that $\beta_1 < 0$: adverse occupational demand shifts slow new work creation.

Estimates of equation (28) in Table 4 confirm this expectation. The first three columns report the relationship between new titles emerging between 1990–2000 and 2000–2018 and occupational exposure to import competition (defined as above, and multiplied by 100) in the same time periods.[42] Panel A uses the inverse hyperbolic sine of new title counts as the dependent variable, whereas panel B uses percentiles of new title counts, relating this to percentiles of import competition. Column 1 presents a basic specification that controls for year main effects, occupational employment shares, and the share of occupational employment across a set of 13 broad industries (one of which is the manufacturing industry). This shows a borderline statistically significant negative effect of demand contractions on new title emergence. This effect becomes significant in column 2 when controlling for occupations' exposure to industry-linked augmentation patents: since manufacturing industries which import from China are relatively innovation-intensive, we underestimate the effect of the domestic industry demand contraction on new title emergence without this control. Because $100 \times$ the import competition measure has a standard deviation of around 1.43, the point

[42] Our estimates are reduced-form models, directly using exposure to Chinese imports by other developed countries. Point estimates are similar when using two-stage least squares (instrumenting exposure to U.S. imports with exposure to other developed countries' imports). The first-stage coefficient at the *industry* level (Borusyak et al. 2021) is 1.20 for imports from the U.S. onto imports from other countries, with a t-statistic of 5.82.

estimate of -0.118 (SE $= 0.051$) means that a one standard deviation greater exposure to import competition predicts a 17 percent reduction in the rate of new title emergence ($-0.118 \times 1.43 \approx -0.17$). Column 3 additionally adds dummies for 13 broad occupation groups. This increases the point estimate to -0.125 and leaves the standard error unaffected. Column 4 further controls for the contemporaneous log change in occupational employment, which tests whether the effect of trade exposure on new title emergence is accounted for by changes in occupational scale. This specification is quite conservative since it arguably controls for an intermediate outcome that is directly affected by the China trade shock. Nevertheless, the point estimate is only slightly smaller in this specification and equal to the column 2 estimate. Notably, the augmentation exposure measure, which is included in all columns except the first, is a robust positive predictor of new title growth across all specifications.

Do these results reflect trends in new work creation that are endemic to U.S. manufacturing and predate the China trade shock? As a check on this possibility, columns 4 – 6 of Table 4 perform a placebo test where we use as the dependent variable new title counts from the twenty years *prior* to the rise of imports from China, 1970–1980 and 1980–1990 ($t = 1980, 1990$), and relate this to import exposure over the entire 1990–2018 period. Consistent with our interpretation of the earlier columns, Chinese import competition exposure has no explanatory power for the emergence of new occupational titles in *subsequently* trade-exposed occupations; in all cases, the point estimates are *positive* and statistically insignificant. This pattern increases our confidence that the effects identified in the first three columns reflect the causal impact of demand contractions on the rate of new work emergence. As a further robustness test, Panel B repeats all of the prior estimates using occupational percentiles of new title counts in place of the IHS measure. While magnitudes should not be compared across these columns since the dependent variables are in different units, the precision of the point estimates is similar across the panels (either slightly higher or slightly lower, depending on the specification).

Alongside confirming a central tenet of our conceptual framework, these results offer an additional substantive implication. Much evidence documents that rising import competition has depressed employment in trade-exposed industries and associated occupations during the last two decades.[43] The findings in Table 4 make clear that these adverse shocks do not merely reduce employment numerically but also depress the emergence of new categories of

[43]See Autor et al. (2016) for a summary of this evidence.

work—new specialties or 'tasks'—that would otherwise have contributed to labor demand. Thus, adverse demand shifts not only yield *less* work in total but also fewer instances of new work. As we show next, positive demand shifts have the opposite effect.

6.2 Demand expansions and new task emergence

As a second source of demand shifts, we follow DellaVigna and Pollet (2007) in exploiting changes in the demographic structure of the U.S. population between 1980 and 2018 to predict movements in industry-level demands, which in turn affect occupation-level demands.[44] For this approach, we use Bureau of Labor Statistics Consumer Expenditure Survey data to obtain predicted consumption across product categories for household members of different ages. We multiply these age-specific coefficients by U.S. Census population data to construct predicted consumption by product based on the evolution of population age structure. We crosswalk these consumption patterns to consistent Census industries to obtain predicted relative demand shifts for consistent industries over 1980–2018.[45] We finally measure occupational exposure to demographically-induced demand shifts, DemandX_{jt}, within 301 consistently defined occupations for 1980–2018 by calculating:

$$\text{DemandX}_{jt} = \sum_i \frac{E_{ij,t-1}}{E_{j,t-1}} \times \tilde{\Delta} \ln \text{demand}_{i,t}. \tag{29}$$

Here $\tilde{\Delta} \ln \text{demand}_{i,t}$ is the predicted log change in demand for output of industry i in time interval t, $E_{ij,t-1}$ is the 20-year lag of employment of occupation j in industry i, and $E_{j,t-1}$ is the 20-year lag of employment of occupation j across all industries.

Reflecting the sharp changes population age structure induced by the aging of the Baby Boom cohorts, Figure 9 shows that occupations' exposure to demographic demand shifts differed substantially across decades.[46] Between 1980 and 2000, when the Baby Boom cohorts were rapidly expanding the prime-aged population, demographic forces raised demand for childcare workers, real estate, and sales-related occupations. In the second period, when the Baby Boom cohorts were entering late working age and retirement, demographic demands

[44] DellaVigna and Pollet (2007) use population aging to predict long-run stock market price changes for industries impacted by demand increases from this source.

[45] These demand measures account for inter-industry input-output linkages, and hence correspond to final demands. Appendix D.2 provides details.

[46] Appendix Figure A2 plots these changes in age structure which drive our subsequent estimation.

shifted towards personal service and health occupations. Education occupations experienced positive demand shifts throughout, driven by growing cohorts of children and young adults in the two sub-periods, respectively.

Following our analysis of the China trade shock, we regress new title emergence on the demographic demand shift measure with the following specification,

$$\text{IHS}\left(\text{newtitles}_{jt}\right) = \beta_1 \text{DemandX}_{jt} + D_t + Z_{jt} + \varepsilon_{jt}, \tag{30}$$

where Z_{jt} is a vector of controls, including occupational exposure to augmentation and overall occupational employment shares. Some specifications additionally control for broad occupation fixed effects and/or occupations' employment shares across 13 broad industry groups. Greater exposure to the demographic demand shock implies *rising* demand (opposite to exposure to the China trade shock). Our conceptual framework predicts that $\beta_1 > 0$: occupations whose outputs are disproportionately demanded by age groups with increasing shares in the population are expected to experience more rapid emergence of new titles. We estimate equation (30) using stacked first differences for 1980–2000 and 2000–2018, matching the frequency of the new title data.

Table 5 shows that positive demand shifts predict the emergence of new occupational titles. In Panel A, we report the relationship between the inverse hyperbolic sine of new occupational titles emerging between 1980–2000 and 2000–2018 and contemporaneous occupational exposure to demographic demand shifts (multiplied by 100 for clarity). Columns 2, 3, 5 and 6 control for occupational augmentation exposure. Models in columns 3 and 6 additionally control for occupations' employment shares across broad industries, and models in columns 4, 5 and 6 further control for broad occupation fixed effects. Across all columns, occupations with higher predicted demand growth stemming from demographic change exhibit faster new title emergence. These estimates are precise even within broad occupational groups (columns 4, 5 and 6). The standard deviation of 100 $\times$ the demand shift measure is around 1 in the first time interval and 1.5 in the second interval, implying that a one standard deviation relative demand shift increases new title emergence by around 21 percent (using the column 5 estimate, $0.171 \times 1.25 \approx 0.214$). In place of the logarithmic specifications in Panel A, Panel B parameterizes the independent and dependent variables in decade-specific percentiles. Demand shifts have a robust relationship with new title emergence in this alternative specification as well, though the units do not in this case have a direct interpretation.

Figure 10 shows that a substantially different set of occupations is exposed to demand shifts than is exposed to augmentation. Using partial predicted effects from the Table 5 models (panel B, column 5), we find that augmentation exposure and demand exposure are uncorrelated on average over 1980–2018. Logically, not all new work creation is directly related to technological forces. Demand shifts can especially help account for the emergence of new titles in lower-paid personal service jobs such as housekeepers, waiters, and food preparation workers. On the other hand, new title emergence in high-tech jobs such as electrical engineers and computer systems analysts are primarily predicted by augmentation exposure. Notably, a subset of occupations, particularly those in healthcare, is exposed to both positive demand shifts and a high rate of augmentation innovations.

7 Augmentation, Automation, and Employment

Our findings establish that task automation and task augmentation are distinct forces that occur concurrently, often in the same occupations, and yet have dissimilar relationships with new title emergence. What does this imply for occupational employment? Our estimates so far do not answer this question since employment could potentially contract in occupations where new titles are emerging or expand in occupations where tasks are automated. our conceptual model makes a clear prediction, however (Proposition 2): Because new task creation is labor-reinstating, it expands employment and wagebill in the sector where it occurs; and, conversely, because task automation is labor-displacing, it erodes employment and wagebill in the sector where it occurs. We test those predictions here, noting that both the theoretical predictions and empirical tests concern *relative* expansion and contraction of occupational employment and wagebills. Aggregate employment, which is fixed and inelastically supplied in our model, is not the object of study.[47]

Figure 11 motivates our analysis of employment by documenting the striking relationship between the emergence of new occupational tasks and net changes in (relative) occupational employment over 1980–2018. The emergence of new occupational tasks is strongly positively associated with occupational employment growth: a 10 higher emergence rate of new titles predicts (relative) occupational employment growth of 3.1 percent in the contempora-

[47] In Acemoglu and Restrepo (2018), task automation has an ambiguous effect on *aggregate* labor demand due to countervailing substitution and scale effects. Our model makes unambiguous predictions for *relative* changes in employment and wagebill across sectors.

neous four decades ($t = 6.2, n = 303$). Occupations such as Computer software developers; Physicians; Vocational and education counselors; and Hairdressers and cosmetologists saw a sizable emergence of new tasks and a large increase in occupational employment. Occupations such as Printing machine operators; Machine feeders and offbearers; and Laundry and dry cleaning workers saw little growth of occupational tasks and a relative decline in employment over the period.

The correlation in Figure 11 is in some sense expected since we have previously established that demand forces that cause occupations to grow also spur the emergence of new occupational tasks. As a further step, and a novel test of our conceptual framework and empirical toolkit, we directly measure the relationship between augmentation exposure, automation exposure, and employment.

7.1 Employment consequences of new task creation and task automation

We assess the relationship between augmentation and automation exposure and net employment changes in full-time equivalent employment by estimating models of the form:[48]

$$\Delta E_{ij,\tau} = \beta_1 \text{AugX}_{ij,\tau} + \beta_2 \text{AutomX}_{j,\tau} + D_i + D_\tau \left(+D_J + \gamma Z_j\right) + \varepsilon_{ij,\tau.} \tag{31}$$

The dependent variable is the Davis-Haltiwanger-Schuh ('DHS', Davis et al. (1998)) employment change in consistent three-digit industry by occupation cells, defined as $\Delta E_{ij,\tau} \equiv \left(E_{ij,t} - E_{ij,t-1}\right) / \left(0.5 \times \left(E_{ij,t} + E_{ij,t-1}\right)\right)$. In contrast to a specification for log employment changes, the DHS transformation allows occupation-industry cells to emerge when employment goes from zero to non-zero over a time period, and similarly, to disappear when cell employment falls to zero. The independent variables of interest are $\text{AugX}_{ij,\tau}$, quantifying exposure to augmentation in industry-by-occupation cells, and $\text{AutomX}_{j,\tau}$, quantifying exposure to automation in occupation cells. To allow consistent measurement of augmentation, automation, and employment, we analyze the 1980–2018 industry-occupation panel, as above. The inclusion of a full set of 206 industry dummy variables in the regression equation means that the coefficients of interest are identified by changes in within-industry

[48]Full-time equivalent employment is equal to annual hours divided by $1,750$, equal to 35 hours per week at 50 weeks per year. We document in a reference table in the Online Appendix that results for log employment changes are highly comparable those for DHS employment changes, despite the exclusion of occupations that have zero start-of-period or end-of-period employment.

occupational employment, holding constant overall industry employment shifts. Standard errors are clustered on industry-by-occupation cells.

Table 6 reports estimates of equation (31) for the relationship between augmentation exposure, automation exposure, and occupational employment growth, showing results from both long-differenced (panel A) and stacked first-differenced (panel B) specifications. Jobs that are more exposed to augmentation experience faster employment growth (column 1), an effect that is most pronounced across jobs within broad occupational categories (column 2). As shown in column 2, each ten percent increase in augmentation exposure predicts 0.24 percent greater occupational employment growth between 1980 and 2018 ($t = 4.10$).[49] Columns 3 and 4 remove the augmentation exposure measure and replace it with the automation exposure measure. Opposite to the case for augmentation, automation exposure predicts statistically significant declines in occupational employment, a relationship that stems from contrasts both within and between broad occupational categories. In column 4, the automation coefficient is estimated at -2.29 ($t = 2.08$), similar in magnitude though opposite in sign to the corresponding coefficient for augmentation.

The most striking results in Table 6 are found when both the augmentation and automation exposure variables are included in columns 5 and 6. Here, each variable has a quantitatively large and precisely estimated predictive relationship with occupational employment: 10 percent greater augmentation exposure predicts 0.35 percent greater employment growth ($t = 6.16$), and 10 percent greater automation exposure predicts 0.29 lower employment growth ($t = 2.70$). Reflecting the negative correlation between AugX and AutomX, these two coefficients are larger (in absolute terms) when estimated jointly rather than when estimated individually. These relationships are equally visible in the stacked first-difference models reported in panel B of the table.[50]

The Table 6 analysis of occupational employment changes echoes that of recent papers by Kogan et al. (2019) and Webb (2020) with one crucial difference. Those papers provide compelling evidence that occupations more exposed to automation patents (Kogan et al.,

[49]Here, we treat the DHS measure as approximating a log change.

[50]Point estimates are smaller in the stacked first-difference in models due in part to a quirk of the IHS transformation. The stacked first-difference models (representing two decade changes) logically contain many more observations with zero patents than the long-difference specifications. Given the outsized role played by zeros in the IHS transformation, the variance of the automation and augmentation exposure measures is much larger in the short panels, while the covariance with DHS employment changes is comparable. The coefficients of interest (and explanatory power) of the stacked first-difference models in Panel B are accordingly attenuated relative to Panel A, though the standardized effect sizes are comparable.

2019) or software patents (Webb, 2020) exhibited relative employment contractions in recent decades. What our analyses uniquely adds to this body of work is evidence of a countervailing augmentation force, also stemming from innovation, that strongly predicts occupational employment growth. This augmentation force provides independent predictive power for employment changes, a relationship that is only strengthened when accounting simultaneously for automation exposure. We regard this as a crucial finding.

We cannot unfortunately estimate a comparable panel analysis for decades prior to 1980 due to incompatibilities in Census occupation and industry classifications across decades. We can however implement a corresponding exercise for the full sample period using an approximation to employment in new work. Following Lin (2011), we approximate employment in new work as the product of two terms: employment in each three-digit Census occupation; and the fraction of titles in the occupation that have emerged in the most recent decade.[51] Using this approximation, we estimate models of the form

$$\text{IHS}\left(E_{ijt} \times \frac{\text{newtitles}_{jt}}{\text{alltitles}_{jt}}\right) = \beta_1 \text{AugX}_{jt} + D_{It}(+D_{it} + D_{Jt}) + \varepsilon_{ijt} \tag{32}$$

Here, the dependent variable is the inverse hyperbolic sine of the occupational new title share multiplied by employment in occupation-industry cells. Because equation (32) is estimated with industry-by-occupation cells, we can also add industry-related controls, including fixed effects for consistently defined broad industry I, for narrow Census industry i, each interacted with year indicators.[52] Some specifications further add fixed effects for broad occupations by year. The independent variable of interest is occupational exposure to augmentation.[53]

Appendix Table A10 presents estimates of equation 32. In all cases, occupations that are more exposed to augmentation exhibit higher employment in new work: an occupation with a 10 percent greater augmentation exposure exhibits 0.05 to 0.10 percent higher employment in new work. This is also true for Census occupations within the same Census industry (column

[51] We do not directly observe employment in new titles in our primary data sets, but write-in occupational data are available for the no-longer-confidential 1940 Census. As described in Appendix A.2, we apply machine learning techniques to the 1940 Census Complete Count file to assign micro-titles to free text occupational write-in data. This exercise validates the exercise pursued here, albeit imperfectly.

[52] As with occupations, we construct thirteen exhaustive and mutually exclusive broad industries which can be consistently compared over 1940–2018.

[53] In this model specification, we use only augmentation measures obtained by direct linkage to the occupation to avoid measuring a spurious correlation with the dependent variable through occupations' employment distribution across industries.

2), within the same broad occupation group (column 3), within the same broad industry and broad occupation groups (column 4), and within both the same Census industry and the same broad occupation group (column 5). Panel B shows that these results are robust to using year-specific percentiles of employment and augmentation patents in place of the IHS measures in Panel A.

7.2 Sectoral wagebills vs. sectoral employment

We take the sectoral labor demand analysis one step further by assessing the relationship between augmentation exposure, automation exposure, and wagebill changes within occupations (i.e., 'sectors' in our model's terminology). Corollary 2 above makes a clear prediction: responding to sectoral augmentation or automation, sectoral wagebills—that is, the *product* of wages and employment in an occupation–should rise or fall equiproportionately with sectoral employment.

Straightforward as this implication is conceptually, testing it presents an empirical challenge. Because employment grows in augmentation-exposed occupations and contracts in automation-exposed occupations (Table 6), observed wages in these occupations are likely to change purely for compositional reasons, even absent any change in composition-constant wages. As shown by Autor and Dorn (2009); Böhm et al. (2020), the average wage (and more broadly, the expected earnings level) of workers tends to rise in occupations as they contract and fall in occupations as they expand. These compositional changes mechanically generate rising wages in declining occupations and falling wages in expanding occupations. Testing whether augmentation or automation differentially affect composition-constant wages (i.e., the wage per efficiency unit of labor) therefore requires accounting for these potential compositional effects.

We proceed in three steps. A first is to estimate cross-sectional log weekly wage regressions in the primary Census and ACS samples to obtain predicted wages:

$$w_{nt} = \alpha_{nt} + \mathbf{S_n}'\boldsymbol{\beta_{1t}} + (\mathbf{S_n} \times A_n)'\,\boldsymbol{\beta_{2t}} + \left(\mathbf{S_n} \times A_n^2\right)'\boldsymbol{\beta_{3t}} + e_{nt}. \tag{33}$$

Here, w_{nt} is the log hourly earnings of worker n, $\mathbf{S_n}$ is a vector of dummies for completed schooling categories, and A_n is years of age. To account flexibly for education-experience profiles, equation (33) includes a quadratic in age fully interacted with the vector of schooling levels. This model is fit separately for each of eight demographic groups (male/female × white/Black/Hispanic/other) in each time period to form a predicted wage for each worker,

$\tilde{w}_{nt}$.

We next collapse predicted and observed log wage levels into wage means within consistent industry-by-occupation cells. Combining these estimates with cell-level employment, we calculate observed ($W_{ij,t}$), predicted ($\widehat{W}_{ij,t}$), and composition-adjusted ($\widetilde{W}_{ij,t}$) wagebills in each industry-occupation cell, where the composition-adjusted wagebill is equal to the observed minus the predicted wage in an industry-occupation cell multiplied by observed cell-level employment.[54]

We finally estimate the relationship between augmentation exposure, automation exposure, and occupational wagebill changes as

$$\Delta Y_{ij,\tau} = \beta_1 \text{AugX}_{ij,\tau} + \beta_2 \text{AutomX}_{j,\tau} + D_\tau + D_i\,(+D_J + \gamma Z_j) + \varepsilon_{ij}. \tag{34}$$

For these estimates, the dependent variable is the DHS change in one of four variables in the industry-occupation cell: employment ($\Delta E_{ij,\tau}$); wagebill ($\Delta W_{ij,\tau}$); expected wagebill minus employment ($\Delta\widehat{W}_{ij,\tau} - \Delta E_{ij,\tau}$); and composition-adjusted wagebill minus employment ($\Delta\widetilde{W}_{ij,\tau} - \Delta E_{ij,\tau}$).

Table 7 presents estimates. To facilitate comparisons between estimated employment and wagebill effects, the first two columns of the table repeat our preferred specifications for employment ($\Delta E_{ij,\tau}$) from columns 5 and 6 of Table 6. The next two columns report analogous estimates for the relationship between augmentation exposure, automation exposure, and wagebills ($\Delta W_{ij,\tau}$). Consistent with expectations, occupations exposed to greater augmentation exhibit differential wagebill growth, and conversely, those exposed to more automation exhibit differential wagebill declines. A comparison of columns 2 and 4 reveals that both the increase in wagebill associated with augmentation and the decrease in wagebill associated with automation are *smaller* in absolute magnitude than the corresponding changes in employment. Taken literally, this pattern would imply that relative wages are falling in augmentation-exposed occupations and rising in automation-exposed occupations.

Columns 5 and 6 clarify the relationship between employment and wagebill changes by regressing the *expected* change in the wagebill in an industry-occupation cell net of em-

[54] An industry-occupation cell with no employment has a zero wagebill. The DHS transformation allows us to calculate the quasi-log change in the wagebill in an industry-occupation cell provided that it has non-zero employment at either the start or end of the period (or both). A small subset of industry-occupation cells have positive employment in the IPUMS samples but no valid wage data, because all workers in the cell are self-employed or report earnings that are out of scope. We impute the predicted wage for these cells using fitted values from equation 33.

ployment change ($\Delta \widehat{W}_{ij,\tau} - \Delta E_{ij,\tau}$) on augmentation and automation measures. Consistent with Böhm et al. (2020), we find that shifts in worker composition in augmentation-exposed cells predict a significant *fall* in cell-level wages, and conversely, compositional shifts in automation-exposed cells predict a significant *rise* in wages. Concretely, augmentation-exposed occupations gain younger, less-experienced, and somewhat less educated workers as they expand, while automation-exposed occupations effectively 'age in place' as they contract (Autor and Dorn, 2009). These compositional shifts, which are akin to quantity rather than price changes in an earnings equation, cloud inference on the earnings of workers of given skill level.

Columns (7) and (8) purge the influence of compositional shifts by taking as its dependent variable the *composition-adjusted* wagebill change net of employment change ($\Delta \widetilde{W}_{ij,\tau} - \Delta E_{ij,\tau}$). This variable measures the 'excess' (i.e., residual) change in the wagebill in an industry-occupation cell net of shifts in worker composition. Consistent with the logic of our model (i.e., Corollary 2), we obtain relatively tightly estimated zeros—finding no statistically or economically significant relationship between augmentation exposure, automation exposure, and excess changes in industry-occupation wagebills. While aggregate wage levels are surely affected by sectoral changes in labor demand spurred by automation and augmentation, these wage changes accrue at the level of detailed skill groups and hence are not captured by cross-sector comparisons that condition on skill. The resulting demands shifts are, however, visible as sectoral shifts in employment as documented in Table 6.[55]

In summary, Tables 6 and 7 confirms a central implication of the task framework that motivates our analysis: despite their positive correlation, augmentation and automation have opposing implications for sectoral labor demand. These opposite-signed relationships are especially notable given that both technology measures are derived from the same underlying corpus of patents. In the terminology of Acemoglu and Restrepo (2019), we find that automation is task-displacing and augmentation is task-reinstating. In conjunction with evidence above that the locus of innovation has shifted across sectors—from middle-educated, production-oriented sectors, such as mining, manufacturing processes, and transportation, to

[55]We have confirmed the robustness of the wagebill relationships in Table 7 to a variety of permutations, including: using log hourly wages of all employees as our wage measure rather than log weekly wages of full-time full-year workers; using percentile measures of augmentation and automation exposure in place of their IHS values (see Appendix Tables A11 and A12); and replacing the augmentation exposure measure with direct counts of new titles. We have also re-estimated our models for occupational employment (Table 6) using data on only the subset of industry-occupation cells where earnings data are also observed. These results are comparable to, though less precise than, the Table 6 estimates.

primarily highly-educated ones, including electricity and electronics, instruments and information, and healthcare—these results imply that a focus on both automation and new work creation can help to illuminate the sources of rising demand for highly-educated workers and the polarization of non-college employment across multiple decades.

8 Conclusion

The majority of jobs performed in 2018 did not exist in 1940. Much recent empirical work has focused on the displacement of labor from existing job tasks through automation, but is mostly silent on the countervailing force of labor reinstatement through the creation of new tasks. Complementing work by Lin (2011); Acemoglu and Restrepo (2018, 2019); Atalay et al. (2020); Deming and Noray (2020), we construct a novel and detailed inventory of new labor-using tasks that allows us to study the evolution and origins of new work over 1940–2018 using representative U.S. Census and ACS data.

We document that, compared to preexisting ('old') work, new work arises disproportionately in middle-paid production and office work in the first four post-War decades. By contrast, new work creation since 1980 has increasingly concentrated in high-education specialties and, to a lesser extent, low-education personal services. We trace the origins of new work creation to both augmentation and demand shifts: the changing locus of new work emergence results from occupations having had different exposures to these over time. Lastly, we directly contrast new task creation from augmentation with task displacement from automation, finding that these twin technological forces, while positively correlated, have strikingly opposing consequences for employment growth.

References

Acemoglu, D. (1998). Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *The Quarterly Journal of Economics*, 113(4):1055–1089.

Acemoglu, D. and Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, 4:1043–1171.

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., and Price, B. (2016). Import Competition and the Great US Employment Sag of the 2000s. *Journal of Labor Economics*, 34(S1):S141–S198.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (December 2020a). AI and Jobs: Evidence from Online Vacancies. Technical report, National Bureau of Economic Research, Working Paper 28257.

Acemoglu, D., Lelarge, C., and Restrepo, P. (2020b). Competing with Robots: Firm-Level Evidence from France. In *AEA Papers and Proceedings*, volume 110, pages 383–88.

Acemoglu, D. and Restrepo, P. (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 Restrepo, P. (2019). Automation and New Tasks: How technology Displaces and Reinstates Labor. *Journal of Economic Perspectives*, 33(2):3–30.

Acemoglu, D. and Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy*, 128(6):2188–2244.

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The Skill Complementarity of Broadband Internet. *The Quarterly Journal of Economics*, 130(4):1781–1824.

Alekseeva, L., Azar, J., Gine, M., Samila, S., and Taska, B. (2020). The Demand for AI Skills in the Labor Market.

Arntz, M., Gregory, T., and Zierahn, U. (2017). Revisiting the Risk of Automation. *Economics Letters*, 159:157–160.

Atack, J., Margo, R. A., and Rhode, P. W. (2019). "Automation" of Manufacturing in the Late Nineteenth Century: The Hand and Machine Labor Study. *Journal of Economic Perspectives*, 33(2):51–70.

Atalay, E., Phongthiengtham, P., Sotelo, S., and Tannenbaum, D. (2020). The Evolution of Work in the United States. *American Economic Journal: Applied Economics*, 12(2):1–34.

Autor, D. and Dorn, D. (2009). This Job is "Getting Old": Measuring Changes in Job Opportunities using Occupational Age Structure. *American Economic Review*, 99(2):45–51.

Autor, D., Goldin, C., and Katz, L. F. (2020). Extending the Race Between Education and Technology. *AEA Papers and Proceedings*, 110:347–51.

Autor, D. H. (2019). Work of the Past, Work of the Future. In *AEA Papers and Proceedings*, volume 109, pages 1–32.

Autor, D. H. and Dorn, D. (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., Dorn, D., and Hanson, G. H. (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review*, 103(6):2121–68.

Autor, D. H., Dorn, D., and Hanson, G. H. (2016). The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade. *Annual Review of Economics*, 8:205–240.

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

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

Autor, D. H., Katz, L. F., and Krueger, A. B. (1998). Computing Inequality: Have Computers Changed The Labor Market? *The Quarterly Journal of Economics*, 113(4):1169–1213.

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

Babina, T., Fedyk, A., He, A. X., and Hodson, J. (2020). Artificial Intelligence, Firm Growth, and Industry Concentration. SSRN Working Paper 3651052.

Balsvik, R., Jensen, S., and Salvanes, K. G. (2015). Made in China, Sold in Norway: Local Labor Market Effects of an Import Shock. *Journal of Public Economics*, 127:137–144.

Bárány, Z. L. and Siegel, C. (2018). Job Polarization and Structural Change. *American Economic Journal: Macroeconomics*, 10(1):57–89.

Berkes, E. (2018). Comprehensive Universe of U.S. Patents (CUSP): Data and Facts. Working paper.

Bernard, A. B., Jensen, J. B., and Schott, P. K. (2006). Survival of the Best Fit: Exposure to Low-Wage Countries and the (Uneven) Growth of US Manufacturing Plants. *Journal of international Economics*, 68(1):219–237.

Böhm, M. J., von Gaudecker, H.-M., and Schran, F. (2020). Occupation Growth, Skill Prices, and Wage Inequality. CESifo Working Paper 7877, Munich.

Bonfiglioli, A., Crinò, R., Fadinger, H., and Gancia, G. (2020). Robot Imports and Firm-level Outcomes.

Borusyak, K., Hull, P., and Jaravel, X. (2021). Quasi-experimental Shift-share Research Designs. *Review of Economics and Statistics*, forthcoming.

Branstetter, L. G., Kovak, B. K., Mauro, J., and Venancio, A. (2019). The China Shock and Employment in Portuguese Firms. National Bureau of Economic Research.

Brynjolfsson, E. and Mitchell, T. (2017). What Can Machine Learning Do? Workforce Implications. *Science*, 358(6370):1530–1534.

Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What Can Machines Learn, and What Does It Mean for Occupations and the Economy? In *AEA Papers and Proceedings*, volume 108, pages 43–47.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and Labor Market dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica*, 87(3):741–835.

Card, D. and Lemieux, T. (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The Quarterly Journal of Economics*, 116(2):705–746.

Chiacchio, F., Petropoulos, G., and Pichler, D. (2018). The Impact of Industrial Robots on EU Employment and Wages: A Local Labour Market Approach. Bruegel working paper.

Chiquiar, D. (2008). Globalization, Regional Wage Differentials and the Stolper–Samuelson Theorem: Evidence from Mexico. *Journal of International Economics*, 74(1):70–93.

Comin, D. A., Danieli, A., and Mestieri, M. (2020). Income-driven Labor Market Polarization. Technical report, National Bureau of Economic Research.

Cortes, G. M., Jaimovich, N., Nekarda, C. J., and Siu, H. E. (2020). The Dynamics of Disappearing Routine Jobs: A Flows Approach. *Labour Economics*, 65:101823.

Dauth, W., Findeisen, S., and Suedekum, J. (2014). The Rise of the East and the Far East: German Labor Markets and Trade Integration. *Journal of the European Economic Association*, 12(6):1643–1675.

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

Davis, S., Haltiwanger, J., and Schuh, S. (1998). *Job Creation and Destruction*, volume 1. The MIT Press, 1 edition.

De Vries, G. J., Gentile, E., Miroudot, S., and Wacker, K. M. (2020). The Rise of Robots and the Fall of Routine Jobs. *Labour Economics*, 66:101885.

DellaVigna, S. and Pollet, J. M. (2007). Demographics and Industry Returns. *American Economic Review*, 97(5):1667–1702.

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

Deming, D. J. and Noray, K. (2020). Earnings Dynamics, Changing Job Skills, and STEM Careers. *The Quarterly Journal of Economics*, 135(4):1965–2005.

Devlin, A., Kovak, B. K., and Morrow, P. M. (2021). The Long-Run Labor Market Effects of the Canada-US Free Trade Agreement. *Carnegie Mellon University Manuscript*.

Dillender, M. and Forsythe, E. (2019). Computerization of White Collar Jobs. Upjohn institute working paper 19-310.

DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach. *Econometrica*, pages 1001–1044.

Dorn, D. (2009). *Essays on Inequality, Spatial Interaction, and the Demand for Skills*. PhD thesis, University of St. Gallen.

Faber, M. (2020). Robots and Reshoring: Evidence from Mexican labor markets. *Journal of International Economics*, 127:103384.

Felten, E., Raj, M., and Seamans, R. C. (2019). The Effect of Artificial Intelligence on Human Labor: An Ability-Based Approach. In *Academy of Management Proceedings*, volume 2019, page 15784. Academy of Management Briarcliff Manor, NY 10510.

Felten, E. W., Raj, M., and Seamans, R. (2018). A Method to Link Advances in Artificial Intelligence to Occupational Abilities. *AEA Papers and Proceedings*, 108:54–57.

Frey, C. B. (2019). *The Technology Trap*. Princeton University Press.

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

Goldin, C. and Katz, L. F. (1998). The Origins of Technology-Skill Complementarity. *The Quarterly Journal of Economics*, 113(3):693–732.

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

Goldin, C. and Margo, R. A. (1992). The Great Compression: The Wage Structure in the United States at Mid-Century. *The Quarterly Journal of Economics*, 107(1):1–34.

Goos, M. and Manning, A. (2007). Lousy and Lovely Jobs: The Rising Polarization of Work in Britain. *Review of Economics and Statistics*, 89(1):118–133.

Goos, M., Manning, A., and Salomons, A. (2009). Job Polarization in Europe. *American Economic Review*, 99(2):58–63.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–2526.

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

Grennan, J. and Michaely, R. (2020). Artificial Intelligence and High-Skilled Work: Evidence from Analysts. *Available at SSRN*.

Griliches, Z. (1981). Market Value, R&D, and Patents. *Economics Letters*, 7(2):183–187.

Haanwinckel, D. (2020). Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution. *UCLA Working Paper*.

Hall, B. H. and Harhoff, D. (2012). Recent Research on the Economics of Patents. *Annu. Rev. Econ.*, 4(1):541–565.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. Technical report, National Bureau of Economic Research.

Harrigan, J., Reshef, A., and Toubal, F. (2021). The March of the Techies: Technology, Trade, and Job Polarization in France, 1994-2007. *Research Policy*.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The Allocation of Talent and US Economic Growth. *Econometrica*, 87(5):1439–1474.

Humlum, A. (2019). Robot Adoption and Labor Market Dynamics. Princeton university working paper.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *the Quarterly journal of Economics*, 108(3):577–598.

Jaravel, X. and Sager, E. (2020). What are the Price Effects of Trade? Evidence from the US and Implications for Quantitative Trade Models.

Katz, L. F. and Autor, D. H. (1999). Changes in the Wage Structure and Earnings Inequality. In *Handbook of Labor Economics*, volume 3, pages 1463–1555. Elsevier.

Katz, L. F. and Murphy, K. M. (1992). Changes in Relative Wages, 1963–1987: Supply and Demand Factors. *The Quarterly Journal of Economics*, 107(1):35–78.

Kelly, B., Papanikolaou, D., Seru, A., and Taddy, M. (2020). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*. Forthcoming.

Kogan, L., Papanikolaou, D., Schmidt, L., and Seegmiller, B. (2019). Technological Change and Occupations over the Long Run. Working Paper 3585676, SSRN.

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

Kovak, B. K. (2013). Regional Effects of Trade Reform: What is the Correct Measure of Liberalization? *American Economic Review*, 103(5):1960–76.

Leonardi, M. (2015). The Effect of Product Demand on Inequality: Evidence from the United States and the United Kingdom. *American Economic Journal: Applied Economics*, 7(3):221–47.

Lin, J. (2011). Technological Adaptation, Cities, and New Work. *Review of Economics and Statistics*, 93(2):554–574.

Mann, K. and Püttmann, L. (2020). Benign Effects of Automation: New Evidence from Patent Texts. SSRN Working Paper 2959584.

Mazzolari, F. and Ragusa, G. (2013). Spillovers from High-Skill Consumption to Low-Skill Labor Markets. *Review of Economics and Statistics*, 95(1):74–86.

Michaels, G., Natraj, A., and Van Reenen, J. (2014). Has ICT Polarized Skill Demand? Evidence from Eleven Countries over Twenty-Five Years. *Review of Economics and Statistics*, 96(1):60–77.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv preprint arXiv:1301.3781*.

Moser, P. (2016). Patents and Innovation in Economic History. *Annual Review of Economics*, 8:241–258.

Pennington, J., Socher, R., and Manning, C. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar. Association for Computational Linguistics.

Pierce, J. R. and Schott, P. K. (2016). The Surprisingly Swift Decline of US Manufacturing Employment. *American Economic Review*, 106(7):1632–62.

Susskind, D. (2020). *A World Without Work: Technology, Automation, and How We Should Respond*. Metropolitan Books, New York, N.Y.

Topalova, P. (2010). Factor Immobility and Regional Impacts of trade Liberalization: Evidence on Poverty from India. *American Economic Journal: Applied Economics*, 2(4):1–41.

U. S. Department of Labor, Employment and Training Administration (1991). *Dictionary of Occupational Titles: Revised Fourth Edition*. U.S. Government Printing Office.

US Census Bureau (1915–2018). *Census of Population: Alphabetical Index of Industries and Occupations*. US Bureau of the Census.

van der Loo, M. (2014). The stringdist Package for Approximate String Matching. *R Journal*, 6(1):111–122.

Webb, M. (2020). The Impact of Artificial Intelligence on the Labor Market. Working paper.

Figure 1: Coding Process of Occupation Write-ins in the ACS

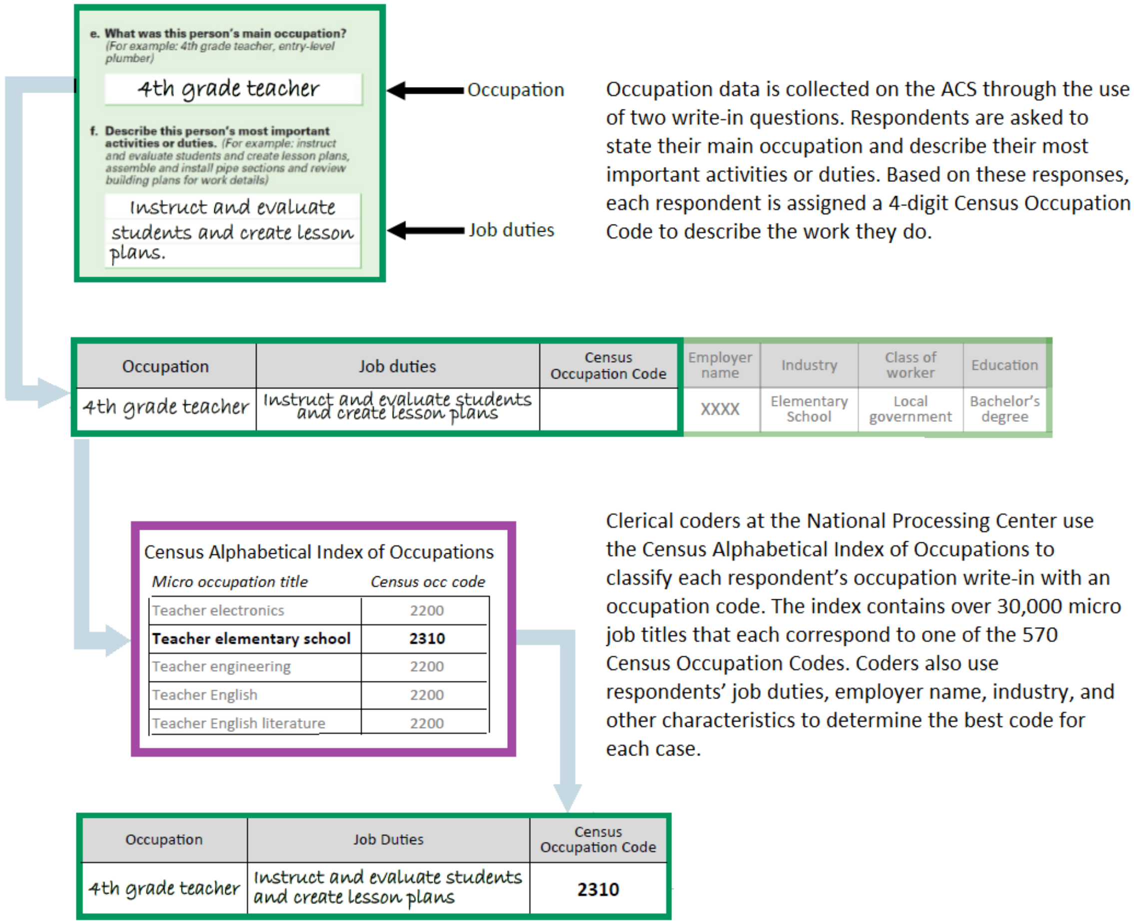


Figure 2: Employment Counts by Broad Occupation in 1940 and 2018, Distinguishing Between Titles Present in 1940 Versus Those Added Subsequently

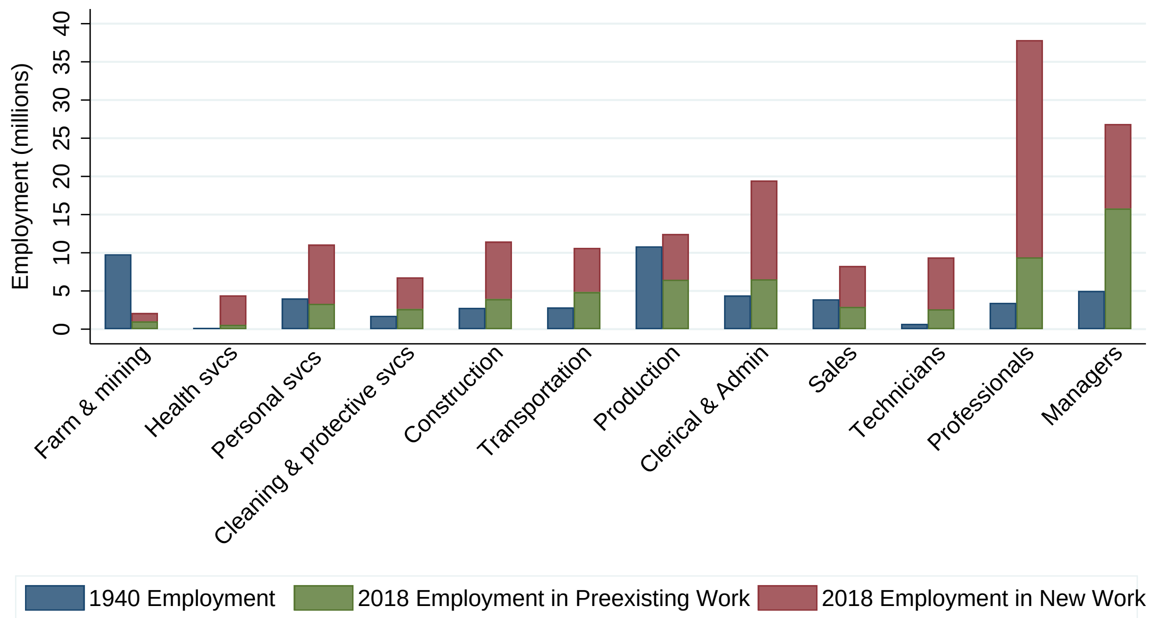


Figure 3: Citation-Weighted Innovation Composition by Broad Technology Class, 1900–2018

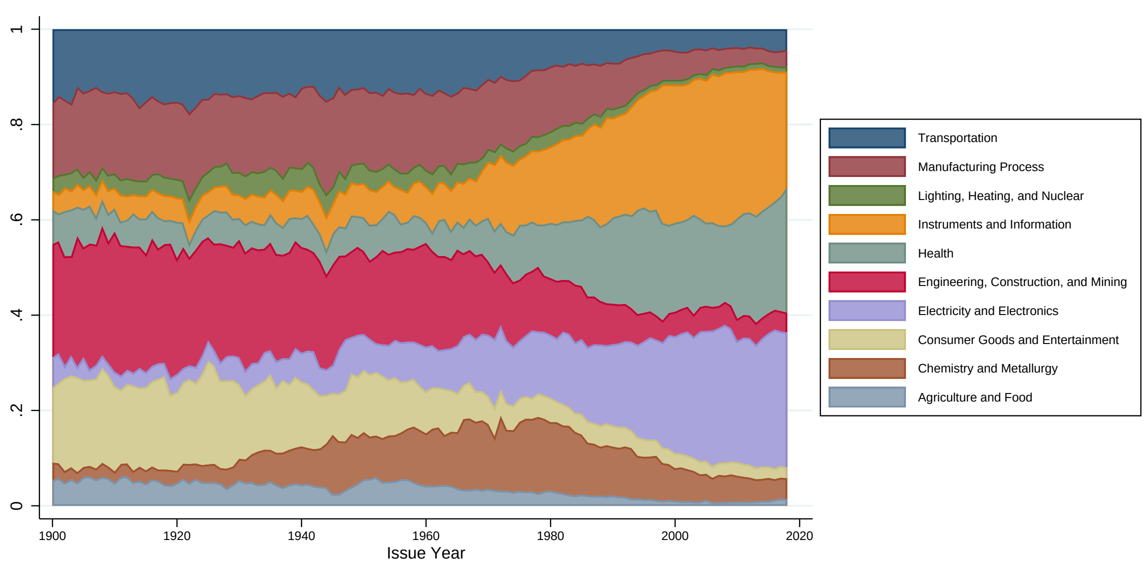


Figure 4: Linking Patents to Occupations and Industries

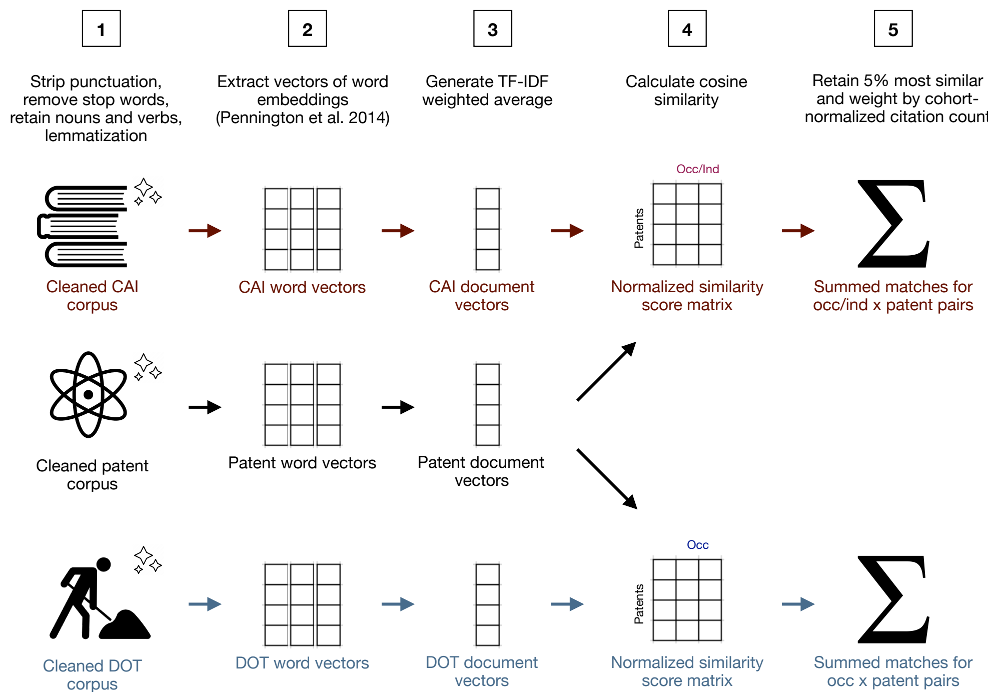
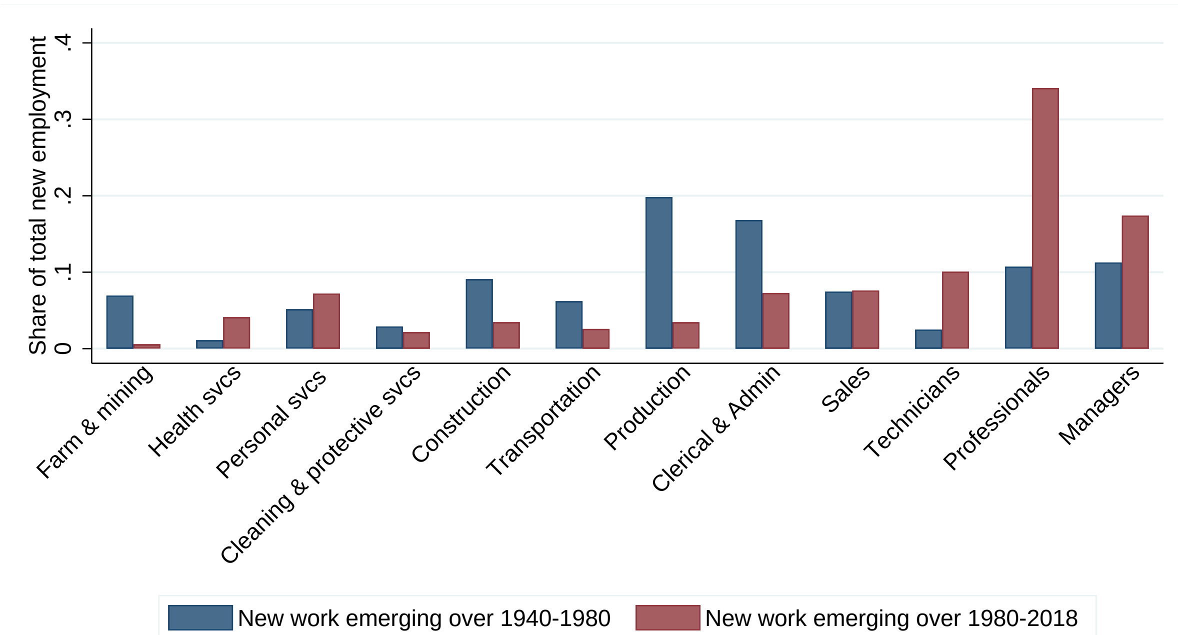


Figure 5: Average Occupational Employment Shares of New Work, 1940–1980 and 1980–2018



Note: Figure shows the average occupational distribution of employment in new work identified in 1940, 1950, 1960, 1970, and 1980; compared to new work identified in 1990, 2000, 2010, and 2018.

Figure 6: Occupational Employment Distribution of New and Preexisting Work by Education Group

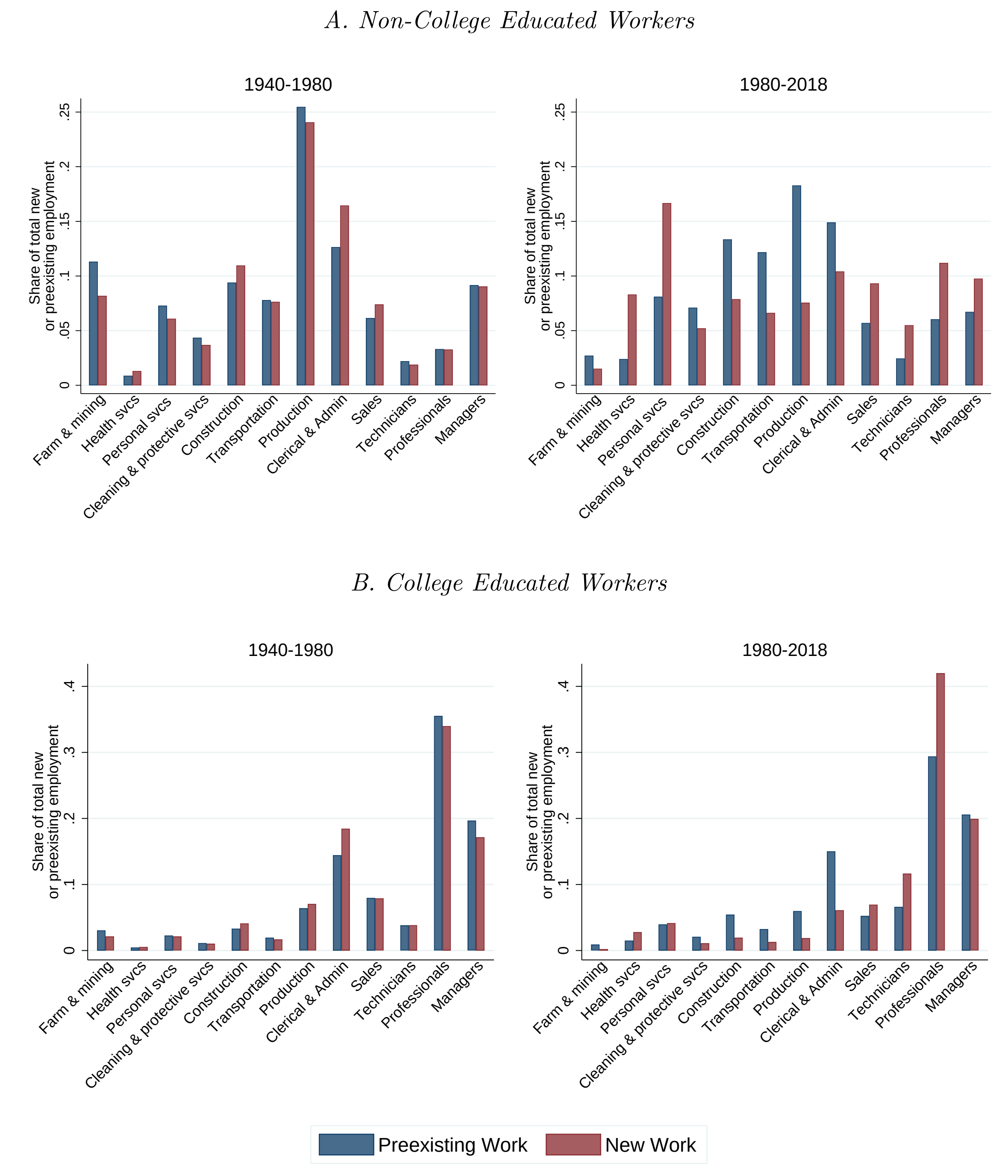
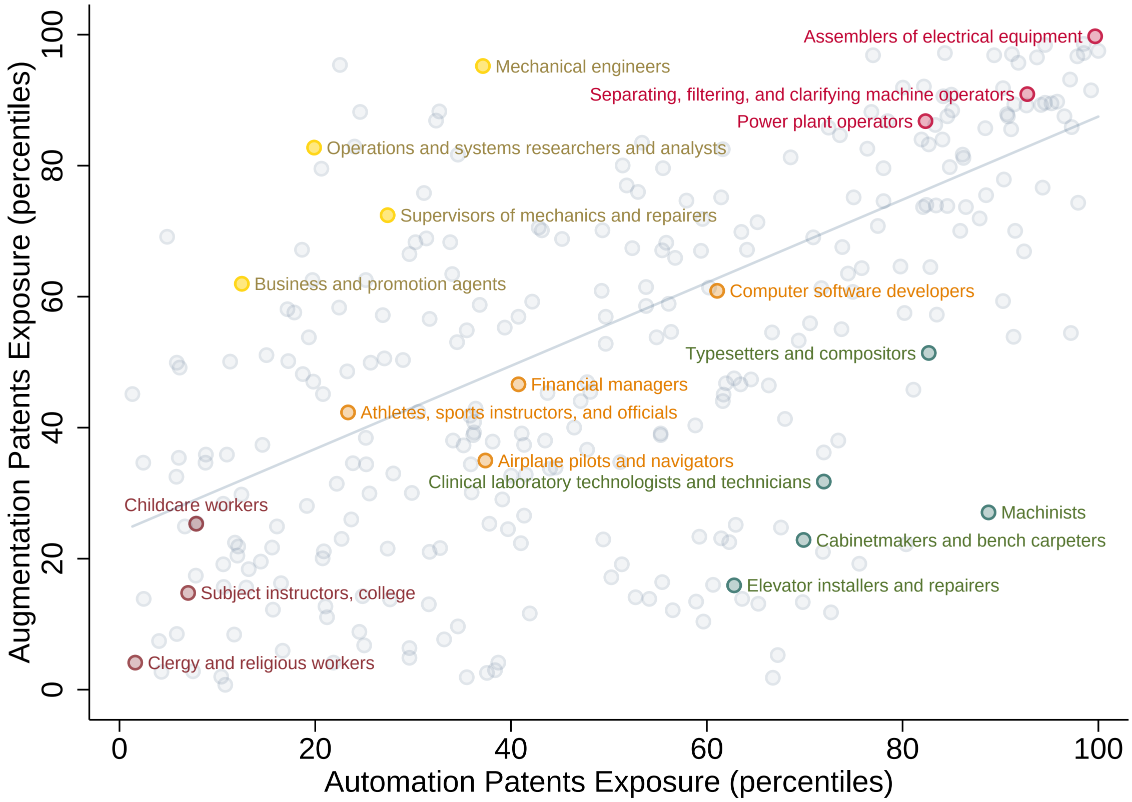
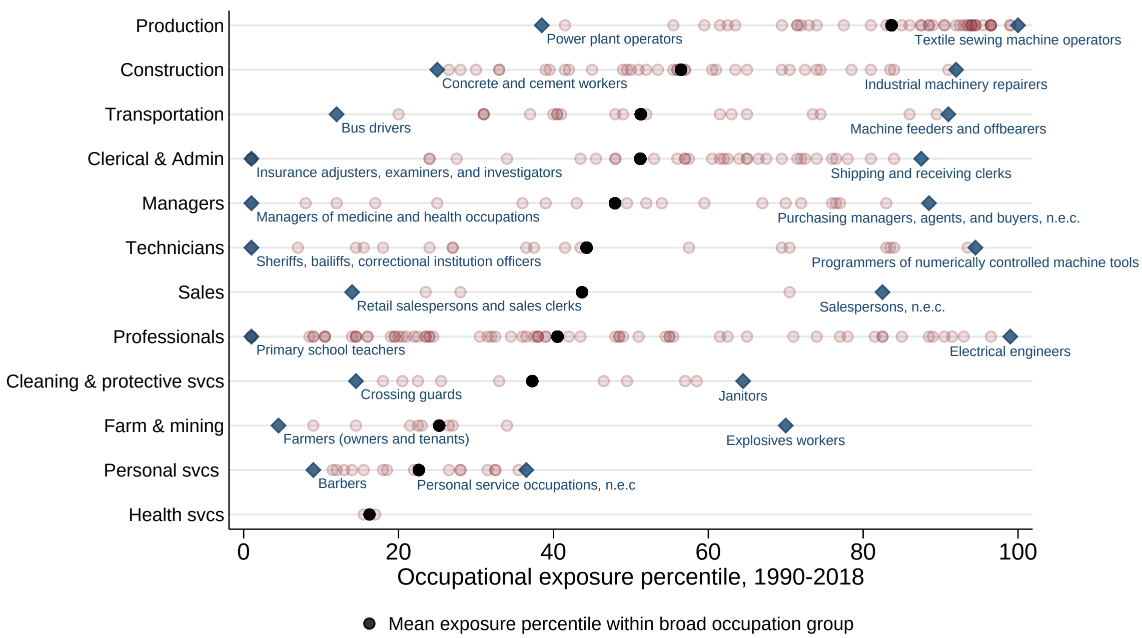


Figure 7: The Relationship between Exposure to Automation and Augmentation Patents at the Occupation Level, 1980–2018



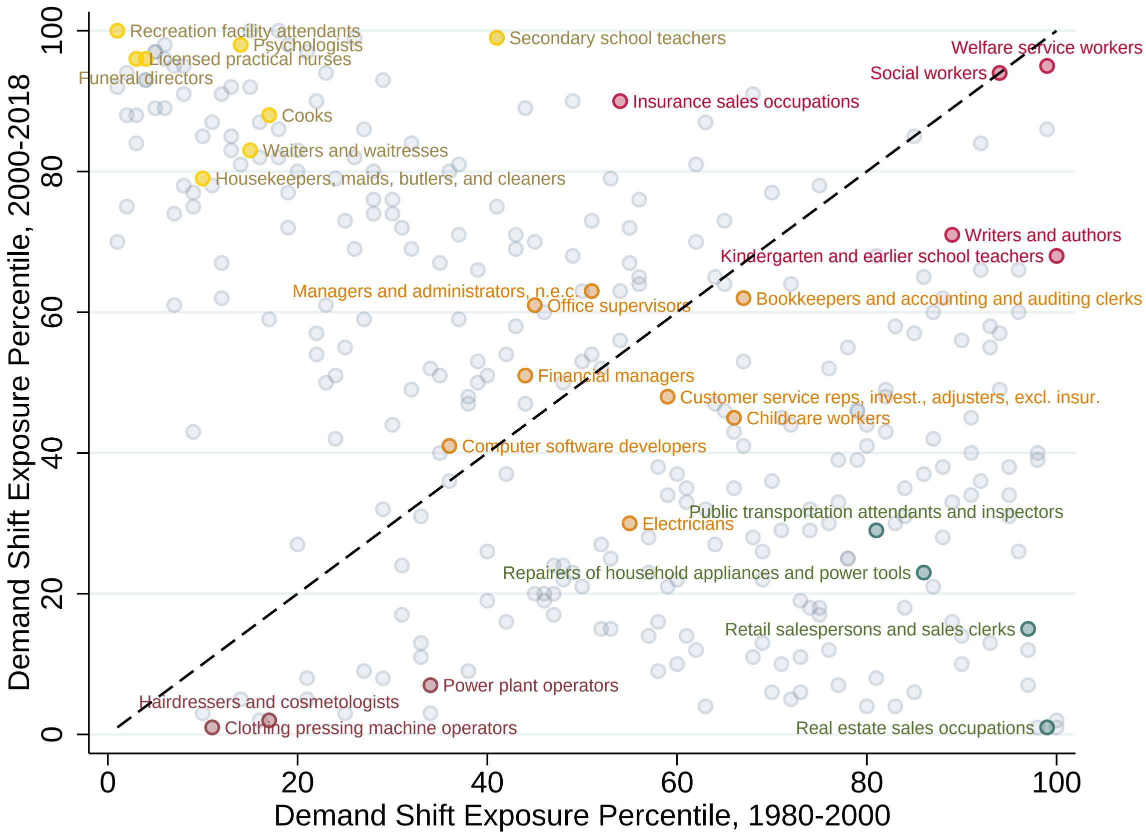
Note: Each point corresponds to the average percentile of automation ($x-$axis) and augmentation ($y-$axis) exposure of one 303 consistently defined three-digit Census occupations ($n = 303$), where the average is taken over 1980, 1990, 2000, 2010, and 2018. Plotted employment-weighted regression line has slope of 0.63 (SE = .076) and intercept of 0.24 with $R^2 = 0.37$.

Figure 8: Percentiles of Occupational Exposure to Import Competition from China, by Broad Occupation



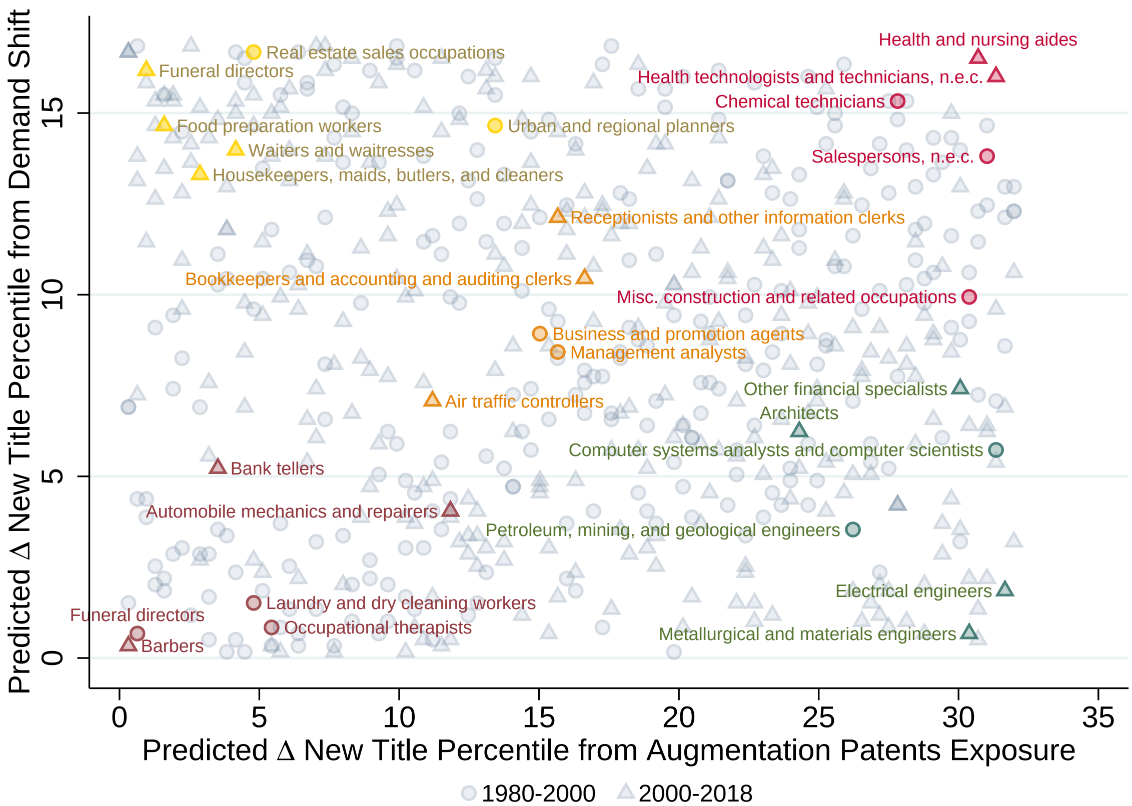
Note: Figure shows percentiles of exposure to import competition (averaged over sub-periods 1990–2000 and 2000–2018) for consistent Census occupations classified into twelve broad occupation groups. Occupation groups are ranked by their average exposure percentile across occupations.

Figure 9: Percentiles of Occupational Exposure to Demand Shifts from Demographic Change, 1980–2000 and 2000–2018



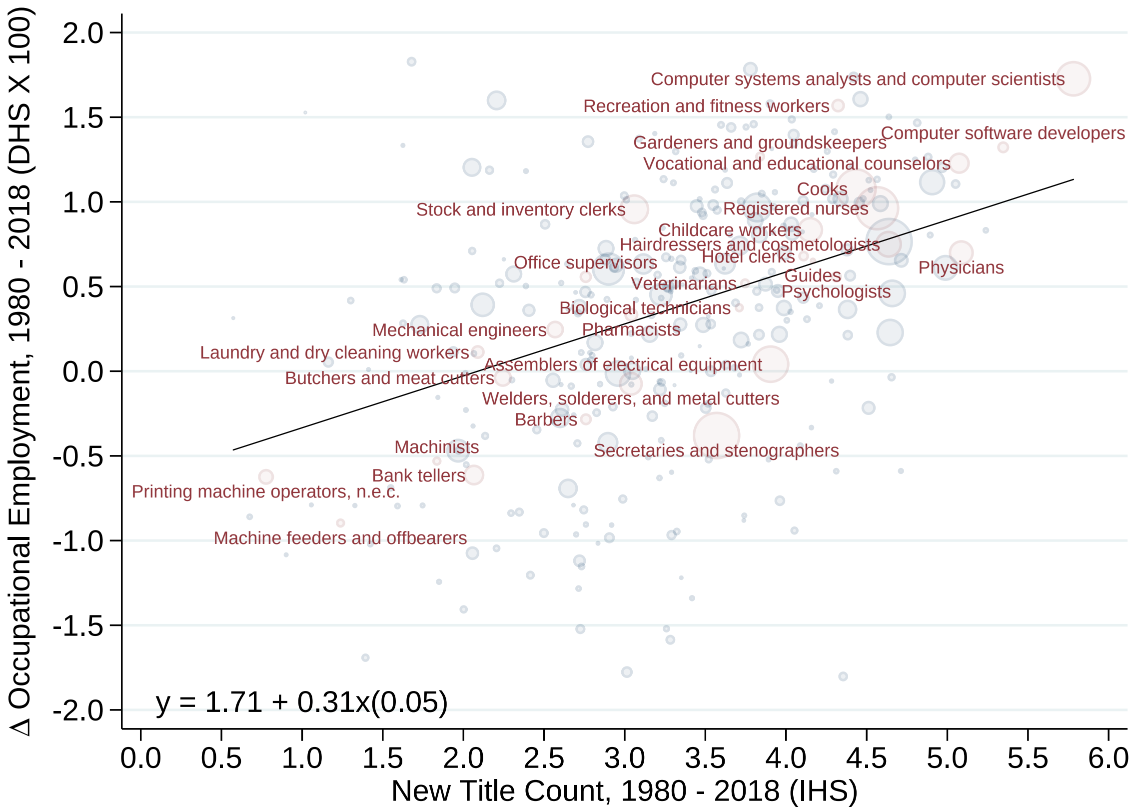
Note: Figure shows percentiles of exposure to demographic demand shifts over 1980–2000 versus 2000–2018 for consistent Census occupations. The dashed line indicates 45 degrees.

Figure 10: Predicted Δ Occupational New Title Share Percentile From Exposure to Augmentation vs. Exposure to Demand Shifts, 1980–2000 and 2000–2018



Note: Figure shows partial predicted new title percentile from exposure to augmentation plotted against partial predicted new title percentile from exposure to demand shifts, for consistent Census occupations. Predictions are based on column 5 in panel B from Table 5.

Figure 11: New Title Count vs. Predicted Δ Occupational Employment, 1980–2018



Note: Figure shows predicted Inverse Hyperbolic Sine (IHS) new occupation title count plotted against predicted Davis Haltiwanger Schuch (DHS) employment growth in consistent Census occupations.

Table 1: Examples of New Titles by Decade, 1940–2018

Year	Example Titles Added between Prior Census and Current Year	
1940	Automatic welding machine operator	Acrobatic dancer
1950	Airplane designer	Tattooer
1960	Textile chemist	Pageants director
1970	Engineer computer application	Mental-health counselor
1980	Controller, remotely-piloted vehicle	Hypnotherapist
1990	Circuit layout designer	Conference planner
2000	Artificial intelligence specialist	Amusement park worker
2010	Technician, wind turbine	Sommelier
2018	Cybersecurity analyst	Drama therapist

Notes: Examples of new 'micro' titles added to Census Alphabetical Index of Occupations by year, 1940–2018.

Table 2: Occupational New Title Emergence and Augmentation Exposure, 1940–2018

	(1)	(2)	(3)	(4)	(5)	(6)
			A. Occupational New Title Count IHS			
AugX	0.159***	0.115***	0.096***			
(Pat Count IHS, Ind-Link)	(0.034)	(0.023)	(0.018)			
AugX				0.130***	0.127***	0.125***
(Pat Count IHS, Occ-Link)				(0.016)	(0.012)	(0.010)
N	3,668	3,668	3,668	3,668	3,668	3,668
R^2	0.634	0.674	0.754	0.679	0.718	0.795
			B. Occupational New Title Count Pctl			
AugX	0.114**	0.119**	0.095**			
(Pat Count Pctl, Ind-Link)	(0.041)	(0.039)	(0.035)			
AugX				0.291***	0.349***	0.360***
(Pat Count Pctl, Occ-Link)				(0.034)	(0.038)	(0.034)
N	3,668	3,668	3,668	3,668	3,668	3,668
R^2	0.331	0.356	0.453	0.376	0.409	0.507
Year FE	X	X		X	X	
Broad Occ FE		X			X	
Broad Occ × Year FE			X			X
Occ Emp Shares	X	X	X	X	X	X

Notes: Census occupations over 1940–2018. Models weighted by annual occupational employment shares. Broad occupations are 12 groups consistently defined over time. Robust standard errors in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 3: Occupational New Title Emergence and Augmentation versus Automation Exposure, 1980–2018

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A. Occupational New Title Count IHS, 1980–2018						
AugX	0.146***	0.158***	0.144***		0.154***	0.159***	0.151***
(Pat Count IHS, Occ-Link)	(0.022)	(0.015)	(0.015)		(0.024)	(0.015)	(0.014)
AutomX				0.105**	-0.016	-0.003	-0.030
(Pat Count IHS)				(0.033)	(0.030)	(0.027)	(0.029)
N	1,212	1,212	1,212	1,212	1,212	1,212	1,212
R^2	0.59	0.66	0.73	0.52	0.59	0.66	0.73
	B. Occupational New Title Count Pctl, 1980–2018						
AugX	0.335***	0.440***	0.439***		0.432***	0.483***	0.474***
(Pat Count Pctl, Occ-Link)	(0.075)	(0.053)	(0.052)		(0.091)	(0.054)	(0.051)
AutomX				0.122+	-0.157+	-0.124+	-0.115
(Pat Count Pctl)				(0.067)	(0.082)	(0.073)	(0.073)
N	1,212	1,212	1,212	1,212	1,212	1,212	1,212
R^2	0.24	0.34	0.45	0.17	0.25	0.34	0.45
Year FE	X	X		X	X	X	
Broad Occ FE		X				X	
Broad Occ X Year FE			X				X
Occ Emp Shares	X	X	X	X	X	X	X

Notes: Consistently defined Census occupations over 1980–2018. Models weighted by annual occupational employment shares. Broad occupations are 12 groups consistently defined over time. Standard errors clustered by occupation in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 4: Occupational New Title Emergence and Demand Contractions from Import Competition

	Years 2000 & 2018				Years 1980 & 1990 (Placebo Test)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				A. Occupational New Title Count IHS				
ImportX 2000 & 2018	-0.089+	-0.118*	-0.125*	-0.118*	0.071	0.091	0.105	0.103
(100 × Δ Imports)	(0.052)	(0.051)	(0.051)	(0.052)	(0.139)	(0.086)	(0.076)	(0.077)
AugX 2000 & 2018		0.270***	0.328***	0.325***				
(Pat Count IHS, Ind-Link)		(0.056)	(0.061)	(0.060)				
AugX 1980 & 1990						0.247**	0.298***	0.296***
(Pat Count IHS, Ind-Link)						(0.080)	(0.050)	(0.051)
N	606	606	606	606	606	606	606	606
R^2	0.368	0.435	0.566	0.568	0.581	0.614	0.658	0.659
				B. Occupational New Title Count Pctl				
ImportX 2000 & 2018	-0.117	-0.398**	-0.210*	-0.192+	0.005	-0.174	0.074	0.063
(Δ Import Pctl)	(0.117)	(0.127)	(0.102)	(0.105)	(0.158)	(0.129)	(0.120)	(0.125)
AugX 2000 & 2018		0.518***	0.633***	0.620***				
(Pat Count Pctl, Ind-Link)		(0.092)	(0.082)	(0.084)				
AugX 1980 & 1990						0.320+	0.306*	0.310*
(Pat Count Pctl, Ind-Link)						(0.164)	(0.125)	(0.126)
N	606	606	606	606	606	606	606	606
R^2	0.264	0.359	0.493	0.495	0.225	0.246	0.292	0.294
Broad Occ FE			X	X			X	X
Δ Log Occ Emp				X				X
Year FE	X	X	X	X	X	X	X	X
Occ Emp Shares	X	X	X	X	X	X	X	X
Broad Ind Emp Shares	X	X	X	X	X	X	X	X

Notes: Consistently defined Census occupations. Models weighted by the average of start- and end-period occupational employment shares. Standard errors clustered by occupation in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 5: Occupational New Title Emergence
and Demand Expansions from Demographic Change

	(1)	(2)	(3)	(4)	(5)	(6)
		A. Occupational New Title Count IHS				
DemandX	0.141*	0.186***	0.146***	0.112*	0.171***	0.148***
(100 × Δ Demand)	(0.057)	(0.043)	(0.043)	(0.055)	(0.041)	(0.042)
AugX		0.140**	0.271***		0.242***	0.342***
(Pat Count IHS, Ind-Link)		(0.049)	(0.049)		(0.053)	(0.053)
N	602	602	602	602	602	602
R^2	0.329	0.364	0.464	0.438	0.511	0.572
		B. Occupational New Title Count Pctl				
DemandX	0.188**	0.185**	0.139*	0.179**	0.166**	0.122*
(Δ Demand Pctl)	(0.063)	(0.061)	(0.063)	(0.060)	(0.055)	(0.058)
AugX		0.122+	0.320***		0.320***	0.498***
(Pat Count Pctl, Ind-Link)		(0.067)	(0.070)		(0.063)	(0.065)
N	602	602	602	602	602	602
R^2	0.222	0.236	0.333	0.343	0.405	0.470
Broad Ind Emp Shares			X			X
Broad Occ FE				X	X	X
Year FE	X	X	X	X	X	X
Occ Emp Shares	X	X	X	X	X	X

Notes: Consistently defined Census occupations, 1980–2000 and 2000–2018. Models weighted by the average of start- and end-period occupational employment shares. Standard errors clustered by occupation in parentheses. $^+p < 0.10$, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 6: Occupational Employment Growth and Augmentation versus Automation Exposure, 1980–2018

Dependent variable: 100 × DHS Employment Change in Occupation-Industry Cells

	(1)	(2)	(3)	(4)	(5)	(6)
	A. 1980–2018 Long Difference					
AugX	2.38***	3.24***			3.68***	3.51***
(Pat Count IHS, Ind×Occ-Link)	(0.58)	(0.57)			(0.57)	(0.57)
AutomX			-7.00***	-2.29*	-7.89***	-2.94**
(Pat Count IHS)			(0.85)	(1.10)	(0.84)	(1.09)
N	42,055	42,055	42,055	42,055	42,055	42,055
R^2	0.43	0.49	0.44	0.48	0.45	0.49
	B. 1980–2000 & 2000–2018 Stacked First Differences					
AugX	1.05**	1.80***			1.88***	1.97***
(Pat Count IHS, Ind×Occ-Link)	(0.33)	(0.30)			(0.31)	(0.30)
AutomX			-3.63***	-1.21*	-4.15***	-1.62**
(Pat Count IHS)			(0.46)	(0.58)	(0.46)	(0.58)
N	81,328	81,328	81,328	81,328	81,328	81,328
R^2	0.34	0.37	0.35	0.37	0.35	0.37
Broad Occ (× Year) FE		X		X		X
Ind (× Year) FE	X	X	X	X	X	X

Notes: Consistently defined Census occupations and industries over 1980–2018. Models weighted by average annual occupation-industry cell employment shares at the start and end of the time period. Broad occupations and industries are 12 and 13 groups consistently defined over time. Standard errors clustered by occupation-by-industry in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 7: Occupational Wagebill Growth and Augmentation versus Automation Exposure, 1980–2018

Dependent variable: 100 × DHS Employment and Wagebill Changes in Occupation-Industry Cells

	Δ Employment		Δ Wagebill		Δ E[Wagebill] - Δ Employment		Δ Adj. Wagebill - Δ Employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					A. 1980–2018 Long Difference			
AugX	3.68***	3.51***	3.11***	3.10***	-0.62***	-0.50***	0.06	0.11
(Pat Count IHS, Ind×Occ-Link)	(0.57)	(0.57)	(0.57)	(0.55)	(0.11)	(0.09)	(0.09)	(0.08)
AutomX	-7.89***	-2.94**	-8.26***	-2.28*	-0.14	0.48**	-0.22	0.22
(Pat Count IHS)	(0.84)	(1.09)	(0.86)	(1.08)	(0.14)	(0.16)	(0.18)	(0.20)
N	42,055	42,055	42,055	42,055	42,055	42,055	42,055	42,055
R^2	0.45	0.49	0.47	0.52	0.40	0.47	0.40	0.45
					B. 1980–2000 & 2000–2018 Stacked First Differences			
AugX	1.88***	1.97***	1.52***	1.68***	-0.34***	-0.29***	-0.02	-0.00
(Pat Count IHS, Ind×Occ-Link)	(0.31)	(0.30)	(0.32)	(0.30)	(0.06)	(0.04)	(0.05)	(0.04)
AutomX	-4.15***	-1.62**	-4.34***	-1.22*	-0.07	0.33***	-0.12	0.07
(Pat Count IHS)	(0.46)	(0.58)	(0.48)	(0.58)	(0.07)	(0.08)	(0.09)	(0.11)
N	81,328	81,328	81,328	81,328	81,328	81,328	81,328	81,328
R^2	0.35	0.37	0.37	0.40	0.44	0.50	0.33	0.37
Broad Occ (× Year) FE		X		X		X		X
Ind (× Year) FE	X	X	X	X	X	X	X	X

Notes: Dependent variable is Davis-Haltiwanger-Schuh (DHS) change in employment (columns 1-2), DHS change in wagebill (columns 3-4), DHS change in expected wagebill net of employment change (columns 5-6), DHS change in composition-adjusted wagebill net of employment change (columns 7-8). Consistently defined Census occupations and industries over 1980–2018. Broad occupations and industries are 12 and 13 groups consistently defined over time. Standard errors clustered by occupation-by-industry in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Appendix

This supplementary appendix contains additional details on data construction as well as several robustness checks on our baseline results.

A Measuring new work

Here, we describe in more detail how we identify new occupation titles; and how total employment in new work is constructed. We also show that occupational new title shares are informative about the occupational distribution of individual-level employment in new work, using Census Complete Count data.

A.1 Procedure for identifying new occupation titles

To extract new work added to the Census Alphabetical Index of Occupations (CAIO) between Census or ACS years $t-1$ and t we use the following steps:

1. Clean titles in both $t-1$ and t by removing capitalization, punctuation, as well as certain common synonyms and decade-specific format changes we identify from inspection of CAIO volumes. This avoids unnecessarily flagging titles as potentially new ("candidate-new") if they are old titles that have been reformatted or reworded in minor or predictable ways.

   (a) Examples of format and wording changes which we discard from is titles like "Accounting Work, Accountant" and "Ad Writer" being added in t when "Accountant" and "Advertising Writer" already exist in $t-1$.

   (b) We also unify variations of titles which contain the same terms either in full or abbreviated form, such as "db" for database, "pt" for physical therapy, "pv" for photovoltaics, and "qc" for quality control.

   (c) Prior to matching, We reduce -man, -person, -work, -er, -or, -ing, -ist etc. titles to the same word base, e.g. "Salesperson", "Salesman" and "Sales work" are changed to "sales"; "Adviser", "Advisor" and "Advising" are degenerated to "advis", and 'Motorist" is degenerated to "motor".

   (d) We clean plural forms, including those ending in "-s" or "-es", and other specific plural forms such as '-ies" when it is a plural of "-y".

   (e) We also discard new gender-specific or gender-neutral versions of existing titles, e.g. we treat the titles "Actor" and "Actress" as one and the same; as we do "Waiter", "Waitress", and "Waitstaff"; and we discard "Chipper Operator" as new because it replaced "Chipperman",

   (f) We discard word order duplicates that are classified to the same Census occupation (e.g. out of "Television Station Manager", and "Manager, Television Station", we retain only one): these occur because at the time of its conception the alphabetical index was used in printed form– multiple word orders were included to save coders time in looking up entries. We retain any title duplicates classified to different industries or occupations, as this may reflect (increasing) emergence of a type of job (an example is the prevalence of IT-related titles across many industries).

(g) Examples of words we automatically denote as synonyms are "auto" and "automobile"; "equipment operator" and "operator"; "sales", "selling", and "sales representative"; "garbage" and "rubbish"; "aide" and "assistant"; "gage" and "gauge".

2. Exact-match and fuzzy-match cleaned occupation titles between CAIO_t to CAIO_{t-1}. We drop all exact title duplicates between $t-1$ and t, disregarding any spacing differences in titles. For the remainder, we retain the three most similar $t-1$ title match for each t title. Specifically:

   (a) For the exact match, we simply match the cleaned titles in t to $t-1$, discard exact matches, and retain the set of unmatched CAIO_t titles as "candidate-new" titles.

   (b) Next, we fuzzy match the CAIO_t candidate-new titles to all CAIO_{t-1} cleaned titles. We use a fuzzy-matching Jaro-Winkler algorithm which matches based on letter-swaps, implemented in R as the package *stringdist* (van der Loo, 2014). This assigns high similarity (i.e. low distance) scores to titles where a low number of single-character transpositions are required to change one word into the other.[56] It also gives higher similarity to strings matching from the beginning up to some specified length: we set the constant scaling factor determining this at the standard value of 0.1. For example, titles which are identical except for a hand-keying error ("Mechanotheraplst" and "Mechanotherapist") receive a high similarity score.

3. Adjudicate remaining unmatched $t+1$ titles ("candidate-new" titles) by classifying them as new or not new, using a combination of automated assignment and careful manual revision. The large majority of candidate-new titles are manually revised, with only around 1,034 automatically assigned. We observe 273,960 total titles over 1940–2018, of which we identify 28,315 as new over the whole period.

Our overarching principle for determining whether a candidate-new title reflects new work is that the title either is a type of job that was entirely nonexistent in the prior period, or reflects some differentiation or specialization of existing work. These latter cases are much more common, and can arise from new or specialized work domains, specialization in educational or professional requirements, or the use of specialized work methods (e.g. by hand or using a machine). On the other hand, candidate-new titles are discarded (i.e.

[56]Note that we have already discarded word order duplicate titles prior to implementing this algorithm.

marked as not new) when they reflect a renaming, reformatting, or generalization of previously existing work. This manual revision process is time-consuming because it requires looking beyond fuzzy-match results and searching the entire $t-1$ index for comparable work.

We implement these principles into the following specific rules for classifying a candidate-new title as new or not new – while not exhaustive, these rules capture commonly occurring cases. A t candidate-new title is:

1. *New* when it is a differentiation of a t title, e.g. "Clinical Psychologist" is new in 1950 as a differentiation of "Psychologist", and "Assembler, Electrical Controls" is new in 1990 as a differentiation of "Assembler, n.s.". This is by far the most commonly occurring type of new title.

2. *New* when it adds specialized work tools to a $t-1$ title, most commonly 'hand' or 'machine'; or specializes operators and set-up operators. E.g. "Bookkeeping Clerk, Machine" is new in 1970 because before only "Bookkeeping Clerk" was listed; and "Drill-Press Set-Up Operator" is new when it is added to "Drill-Press Operator".

3. *New* when it adds some additional educational or professional differentiation to a $t-1$ title. E.g. "Licensed Addiction Counselor" is new in 2018 as an addition to "Addiction Counselor"; and "Health Therapist, Less Than Associate Degree" is new in 1990 as an addition to "Health Therapist". This is a type of new title that occurs relatively infrequently.

4. *New* when it adds "not specified" or "not elsewhere classified" to a $t-1$ title. This reflects more types of this title are emerging which (for the time being) are listed as n.s. / n.e.c. For example, "Mechanic, Instrument, n. s." is added in 1980.

5. *New* when it bifurcates a $t-1$ title into two separate types, usually marked with "incl", "exc", or "any other". E.g. in 1980 the title "Sitter, exc. Child Care" was new since before only "Sitter" had existed.

6. *Not new* when it simply reorganizes information across various columns of the index for the same title. E.g. "Apprentice Dentist" was discarded as new in 1940 because it already existed in 1930. This is a common reason for discarding candidate-new titles.

7. *Not new* when it is generalization from previously-specified title, e.g. "Ad Taker" is not new in 1980 because it simply subsumes the 1970 titles "Classified-Ad Taker" and "Telephone-Ad Taker"; and "Inspector Agricultural commodities" is not new in 1980 because it subsumes "Inspector Fruit", "Inspector Food", and "Inspector Livestock".

8. *Not new* when it is the same as a $t-1$ title except for filler words. E.g. "Software Applications Developer" is not new in 2018 because the title "Software Developer" already existed before.

9. *Not new* when a title is a combination of two existing titles. E.g. "Inker and opaquer" is not new in 1980 because both "Inker" and "Opaquer" already existed in 1970.

A.2 Using new title shares as a measure of employment in new work

A.2.1 Constructing total employment in new work in 2018

We sum the number of new titles added over 1940–2018, nr_{new}, and divide this by the total number titles in the 2018 index adjusted for titles that were removed $\tilde{nr}_{2018}$, separately by broad occupation J. The adjustment in the total title count consists of adding in the implied total number of removed titles nr_{dead}, if this number is positive. That is, the cumulate new title share over 1940–2018 is $\frac{nr_{new}}{\tilde{nr}_{2018}}$ where $\tilde{nr}_{2018} \equiv nr_{2018} + nr_{dead} \equiv nr_{2018} + nr_{new} - (nr_{2018} - nr_{1940})$.

A.2.2 Comparison of occupational new title shares to individual-level employment in new work in 1940

Since we do not observe individual workers employed in new and old micro occupation titles, we use occupational new title shares as a measure of employment in new work.

Importantly, our analyses relating new work emergence (as well augmentation and automation exposure) to occupational employment growth do not require making any assumptions about employment numbers in new versus old titles. However, it is valuable to explore the relationship between new titles and employment in new work to contextualize our findings.

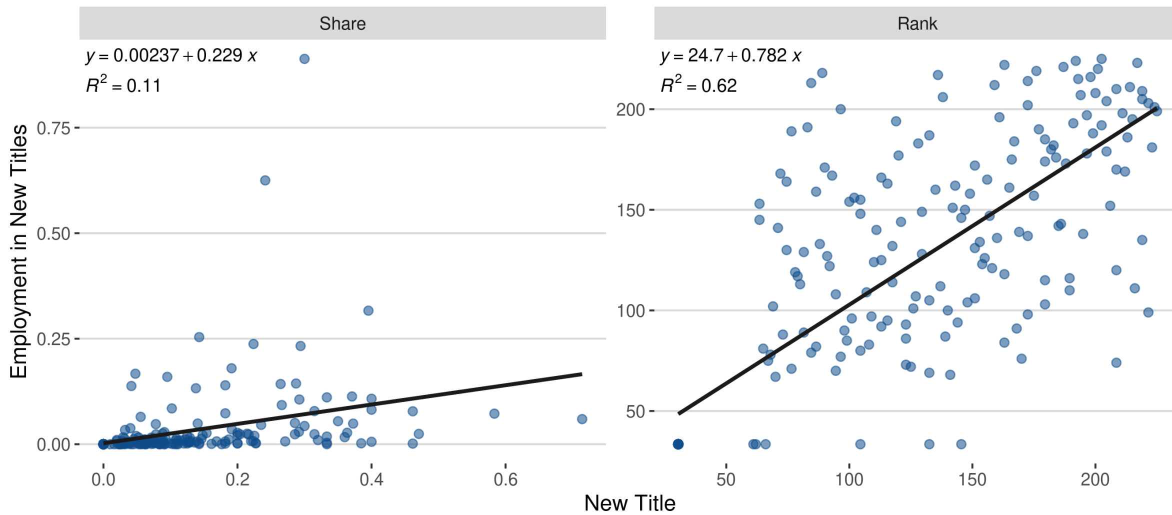
We use individual-level data from the 1940 Census Complete Count (CCC) to compare new title shares and employment in new work. We link 84% of employed, working age individuals in the CCC to micro titles in the Census Alphabetical Index by implementing a combination of fuzzy-matching and term-frequency-inverse-document-frequency (TF-IDF) techniques. Overall, 81% of micro titles in the Census Alphabetical Index are linked to at least one CCC worker.

Individual employment counts in matched occupation titles are aggregated to the 'macro' occupation level to calculate the share of workers employed in each occupation. We also rank occupations by both the share of new titles and the share of employment in new titles to explore relationships between relative measures of new title share and new work share. Figure A1 shows the correlation between new title share and employment share in new titles in the left panel, and the *rank* of new titles and *rank* of employment share in new titles in the right panel.

We find that the relationship between new title share and employment share in new titles is highly significant with a p-value of less than 0.01 for both the regression of new title shares and rank of new title shares. Figure A1 suggests that the rank of new title shares is a better measure for capturing relationships between new title shares and, employment shares in new titles, than the underlying new title shares themselves: the coefficient on new title share and R^2 value increases from 0.229 and 0.115 in regressions using new title shares to 0.782 and 0.615 in regressions using ranks of new title shares.

While the slope in the left panel suggests that employment in new titles is substantially

Figure A1: Comparison of New Title Shares and Employment in New Work, 1940



Note: The left panel shows the relationship between the share of new titles by occupation and the share of employment in new titles. The right panel compares the rank of new title share and the rank of employment across macro occupations, where the lowest rank represents the occupation with the lowest share.

lower than employment in existing titles, there are at least two sources of bias that can generate an underestimate of the true employment count in new titles. First, new titles often represent the specialization of an existing title. Census respondents who are employed in specialized fields may choose to report the more general version of their occupation, while workers with broad occupational responsibilities are unlikely to report specialized occupations. This can cause new specialized occupations to be systematically categorized as existing work, thereby underestimating our estimate of employment in new work. Second, new titles appear to be more difficult to link to census write-ins than existing titles. We find that respondents with CCC write-ins that match exactly to titles in the Census Alphabetical Index have lower rates of employment in new titles compared with workers matched to the Census Alphabetical Index using more flexible matching procedures. As a result, we expect the true rates of employment in new titles to be larger than the employment shares reported in Figure A1.

A.2.3 Characteristics of workers employed in new work and existing work in 1940

Among workers who matched to occupation titles from the Census Alphabetical Index, 1.4% are employed in new titles. In Table A1 we compare the the earnings and education levels of workers employed in new work compared to those employed in pre-existing work. Columns 5 shows that, controlling for age, sex, race, geography, and occupation, workers who earned

\$1000 more in 1940 are 0.1 percentage points more likely to be employed in new work, corresponding to a 7% increase (= 0.1/1.4). Similarly, the likelihood of employment in new work increases substantially with the education level: high school graduates and college graduates are 14% and 36% more likely to be employed in new work relative to a worker with a less than 9th grade education.

We report the most common new titles by education level in Table A2. While some occupations, such as "foreman" and "driver salesman" are accessible across all education levels, new titles that require advanced certifications, such as "petroleum engineer", or "patent attorney" are limited to those with college degrees.

Table A1: Earnings and Education Level for Workers in New vs. Pre-Existing Titles

	Dependent variable: Dummy for being employed in new work				
	(1)	(2)	(3)	(4)	(5)
Earnings	0.002***	0.001***			0.001***
	(0.00001)	(0.00002)			(0.00002)
Education level (Reference category: Less than 9th grade education)					
Some high school			0.005***	0.002***	0.002***
			(0.0001)	(0.0001)	(0.0001)
High school			0.002***	0.003***	0.003***
			(0.0001)	(0.0001)	(0.0001)
Some college			0.003***	0.005***	0.005***
			(0.0001)	(0.0001)	(0.0001)
College			0.001***	0.006***	0.005***
			(0.0001)	(0.0001)	(0.0001)
Occupation FE		X		X	X
Full Controls					X
N	28,660,196	28,660,196	28,143,887	28,143,887	28,143,887
R^2	0.001	0.130	0.0003	0.130	0.130

Notes: Linear probability models. Education estimates compare the probability of employment in new work with workers who have a less than 9th grade education level. Columns 3, 4, and 5 include only observations for workers who are $\geq$ 25 years old. Column 5 includes controls for occupation, age, sex, race, state, and urban/rural status. Earnings measured in thousands in 1940 dollars.
$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table A2: Most Common New Titles by Education Level

Rank	Less Than 9th Grade	High School	At Least Some College
1	c.c.c. foreman	driver salesman	druggist pharmacist
2	driver salesman	c.c.c. foreman	c.c.c. foreman
3	pattern maker	letterer carrier	driver salesman
4	letterer carrier	pattern maker	job interviewer
5	metal finisher	accounting clerk	petroleum engineer
6	route salesman	recreation attendant	naval official
7	c.c.c. worker	druggist pharmacist	accounting clerk
8	share cropper	route salesman	research work or worker
9	spot welder	nurse aid	patent attorney
10	grader operator	helper chemist	research clerk

B Patents

B.1 Linking patents to occupations and industries

Table A3 shows examples of linked patents for a number of Census industries; similarly, Table A4 shows of examples of linked patents for a number of Census occupations.

Table A5 shows the most and least augmentation-exposed consistently defined Census occupations within each broad occupation group, averaged over 1980–2018, where patents have been linked to industry titles. Tables A6 and A8 show the analogous information for occupation-linked augmentation exposure, and automation exposure, respectively.

Table A3: Examples of Individual Patents Linked to Census Industries

Patent name	Linked Industry
A. Examples of Industry Linkages for 1940	
Method of cultivating plants	Agriculture
Variable speed transmission mechanism	Automobile storage, rental, and repair services
Process of making and packaging pie crust dough	Bakery products
Bank-check post-card folder	Banking and other finance
Chewing gum confection	Confectionery
Gypsum lumber	Cut-stone and stone products
Apparatus for pasteurizing liquids	Dairy products stores and milk retailing
Compartment dish for hors d'oeuvre	Eating and drinking places
Wirelessly energized electrical appliance	Electric light and power
Protective covering for plants	Forestry except logging
Collapsible table and seat	Furniture and store fixtures
Knitted stocking foot protector	Knit goods
Medicine glass cover and marker	Medical and other health services
Process of making abrasive materials	Miscellaneous chemical industries
Stapling device	Miscellaneous iron and steel industries
Manufacture of color-pictures	Paints, varnishes, and colors
Radio amplifying system	Radio broadcasting and television
Method of making viewing gratings for relief or stereoscopic pictures	Theaters and motion pictures
B. Examples of Industry Linkages for 2018	
Mobile phone payment processing methods and systems	Accounting, tax preparation, bookkeeping, and payroll services
Methods and compositions for improving plant health	Agricultural chemical manufacturing
Portable folding type hairstyling tool	Beauty salons
Digital video capture system and method with customizable graphical overlay	Computer and peripheral equipment manufacturing
Cropping systems for managing weeds	Crop production
Devices and methods for treating pain associated with tonsillectomies	Drugs, sundries, chemical and allied products wholesalers
Multi-pronged spear-fishing spear tip	Fishing, hunting and trapping
Showerhead mounting to accommodate thermal expansion	Household appliance manufacturing
System and method for data publication through web pages	Newspaper publishers
Video visitation system and method for a health care location	Offices of other health practitioners
Hemoglobin display and patient treatment	Outpatient care centers
Abuse-proofed dosage system	Pharmacies and drug stores
Suspendable and stackable vertical planter	Pottery, ceramics, and plumbing fixture manufacturing
Method and system for securing online identities	Securities, commodities, funds, trusts, and other financial investments
Golf club head with adjustable center of gravity	Sporting goods, and hobby and toy stores
Document revisions in a collaborative computing environment	Software publishers
Controlling power consumption of a mobile device based on gesture recognition	telecommunications, except wired telecommunications carriers
Wired, wireless, infrared, and powerline audio entertainment systems	Wired telecommunications carriers

Table A4: Examples of Individual Patents Linked to Census Occupations

Patent name	Linked Occupation
A. Examples of Occupation Linkages for 1940	
Thermal insulating material	Asbestos and insulation workers
Pie making process	Bakers
Towing dolly	Chauffeurs and drivers, bus, taxi, truck, and tractor
Process for the production of antiseptic agents	Chemical engineers
Corn popper	Cooks–except private family
Lever locking device	Cranemen, hoistmen, and construction machinery operators
Toecap for toe dancing shoes	Dancers, dancing teachers, and chorus girls
Spring roller venetian blind	Decorators and window dressers
Multiple elevator system	Elevator operators
Artificial fish bait	Fishermen and oystermen
Fruit squeezer	Fruit and vegetable graders and packers–except in cannery
Combination hand weeder and cultivator	Gardeners–except farm and groundskeepers
Mail covering	Mail carriers
Variable speed power transmission mechanism	Mechanics and repairmen–railroad and car shop
Chord finder for tenor banjos	Musicians and music teachers
Roof sump or floor drain	Plumbers and gas and steam fitters
Guided transmission of ultra high frequency waves	Radio and wireless operators
Telephone and telegraph signaling system	Telegraph operators
B. Examples of Occupation Linkages for 2018	
Systems and methods for unmanned aerial vehicle navigation	Aircraft pilots and flight engineers
Stabilised supersaturated solids of lipophilic drugs	Chemists and materials scientists
Telepresence robot with a camera boom	Communications equipment operators, all other
Systems and methods for detecting malware on mobile platforms	Computer programmers
Method of treating Attention Deficit Hyper-Activity Disorder	Counselors, all other
Document revisions in a collaborative computing environment	Editors
Mobile personal fitness training	Exercise trainers and group fitness instructors
Broccoli based nutritional supplements	Food cooking machine operators and tenders
Insulation with mixture of fiberglass and cellulose	Insulation workers
Determining text to speech pronunciation based on an utterance from a user	Interpreters and translators
Rotary drill bit including polycrystalline diamond cutting elements	Jewelers and precious stone and metal workers
Method and system for navigating a robotic garden tool	Landscaping and groundskeeping workers
Cuticle oil dispensing pen with ceramic stone	Manicurists and pedicurists
Adaptive audio conferencing based on participant location	Meeting, convention, and event planners
Identification and ranking of news stories of interest	News analysts, reporters, and journalists
Fumigation apparatus	Pest control workers
Low profile prosthetic foot	Podiatrists
Invertible trimmer line spool for a vegetation trimmer apparatus	Tree trimmers and pruners

Table A5: Top and Bottom Augmentation-Exposed Occupations by Broad Occupation Group, Using Ind-Linked Augmentation Patents

Broad Occupation	Consistent Census Occupation	Avg AugX Pctl
Farm & Mining	Explosives workers	51.8
	Farmers (owners and tenants)	9.8
Health Services	Health and nursing aides	29.7
	Dental Assistants	5.8
Personal Services	Personal service occupations, n.e.c	52
	Barbers	4
Cleaning & Protective Services	Guards and police, except public service	71
	Pest control occupations	15.5
Construction	Repairers of data processing equipment	94
	Plasterers	31.5
Transportation	Packers and packagers by hand	85.3
	Bus drivers	13.8
Production	Assemblers of electrical equipment	99.5
	Bakers	19.5
Clerical & Admin	Shipping and receiving clerks	91.5
	Hotel clerks	1.5
Sales	Salespersons, n.e.c.	94.2
	Cashiers	34
Technicians	Engineering technicians	98.3
	Sheriffs, bailiffs, correctional institution officers	4.3
Professionals	Electrical engineers	99.8
	Primary school teachers	1
Managers	Purchasing managers, agents, and buyers, n.e.c.	94
	Funeral directors	5.8

Note: Table shows consistent Census occupations with the highest and lowest augmentation exposure percentile within each broad occupation group, averaged over 1980–2018. Augmentation exposure constructed through industry linkages.

Table A6: Top and Bottom Augmentation-Exposed Occupations by Broad Occupation Group, Using Occ-Linked Augmentation Patents

Broad Occupation	Consistent Census Occupation	Avg AugX Pctl
Farm & Mining	Other mining occupations	83.5
	Fishers, marine life cultivators, hunters, and kindred	36.8
Health Services	Health and nursing aides	59.5
	Dental Assistants	10.5
Personal Services	Recreation facility attendants	80
	Bartenders	1
Cleaning & Protective Services	Laundry and dry cleaning workers	86.5
	Supervisors of cleaning and building service	5.5
Construction	Misc. construction and related occupations	95.3
	Plasterers	4.5
Transportation	Machine feeders and offbearers	97.3
	Helpers, surveyors	19.8
Production	Extruding and forming machine operators	100
	Machinists	5.5
Clerical & Admin	Office machine operators, n.e.c.	90.5
	Proofreaders	4.8
Sales	Cashiers	63.8
	Sales demonstrators, promoters, and models	22.8
Technicians	Engineering technicians	94.8
	Drafters	1.8
Professionals	Chemical engineers	90.8
	Dieticians and nutritionists	1
Managers	Management analysts	67
	Funeral directors	1

Note: Table shows consistent Census occupations with the highest and lowest augmentation exposure percentile within each broad occupation group, averaged over 1980–2018. Augmentation exposure constructed through occupation linkages.

Table A7: Top and Bottom Augmentation-Exposed Occupations by Broad Occupation Group, Using Ind-Occ-Linked Augmentation Patents

Broad Occupation	Consistent Census Occupation	Avg AugX Pctl
Farm & Mining	Other mining occupations	86.5
	Fishers, marine life cultivators, hunters, and kindred	57.2
Health Services	Health and nursing aides	64.9
	Dental Assistants	37.8
Personal Services	Recreation facility attendants	81.1
	Bartenders	25.2
Cleaning & Protective Services	Laundry and dry cleaning workers	70.3
	Housekeepers, maids, butlers, and cleaners	29.1
Construction	Misc. construction and related occupations	95
	Plasterers	37.8
Transportation	Construction laborers	96.7
	Taxi cab drivers and chauffeurs	50.5
Production	Molders and casting machine operators	96.7
	Boilermakers	42.3
Clerical & Admin	Office machine operators, n.e.c.	86.4
	Hotel clerks	29.7
Sales	Salespersons, n.e.c.	82.5
	Sales demonstrators, promoters, and models	58.2
Technicians	Engineering technicians	92
	Dental hygienists	25.4
Professionals	Chemical engineers	93.6
	Physicists and astronomists	12
Managers	Management analysts	81.6
	Funeral directors	8.3

Note: Table shows consistent Census occupations with the highest and lowest augmentation exposure percentile within each broad occupation group, averaged over 1980–2018. Augmentation exposure constructed through occupation-industry linkages.

Table A8: Top and Bottom Automation-Exposed Occupations by Broad Occupation Group

Broad Occupation	Consistent Census Occupation	Avg AutomX Pctl
Farm & Mining	Drillers of oil wells	83
	Fishers, marine life cultivators, hunters, and kindred	32.8
Health Services	Health and nursing aides	55.5
	Dental Assistants	38.8
Personal Services	Recreation facility attendants	45.3
	Guides	5.3
Cleaning & Protective Services	Laundry and dry cleaning workers	61.5
	Supervisors of guards	5.5
Construction	Machinery maintenance occupations	98.3
	Supervisors of mechanics and repairers	27.5
Transportation	Construction laborers	91.5
	Parking lot attendants	12.8
Production	Production checkers, graders, and sorters in manufacturing	100
	Other woodworking machine operators	33
Clerical & Admin	Technicians, n.e.c.	77.8
	Hotel clerks	4.3
Sales	Cashiers	46.8
	Sales demonstrators, promoters, and models	15.3
Technicians	Chemical technicians	94.5
	Sheriffs, bailiffs, correctional institution officers	2
Professionals	Chemical engineers	94.8
	Sales engineers	1
Managers	Purchasing agents and buyers of farm products	47.3
	Managers of medicine and health occupations	10

Note: Table shows consistent Census occupations with the highest and lowest automation exposure percentile within each broad occupation group, averaged over 1980–2018.

C Relationship between patent linkages and task content

In Table A9, we explore the cross-sectional relationship between automation and augmentation in an occupation and the task content of the occupation by regressing each task content measure on patent counts. The measures of "eye-hand-foot coordination", "finger dexterity", "direction, control and planning", "sets limits, tolerances, and standards" and "GED math" task content are obtained from Autor et al. (2003). Following Autor and Dorn (2013), "routine" task content averages "set limits, tolerances, and standards" and "finger dexterity"; and "abstract" task content averages "direction, control and planning" and "GED math". Note that "manual" task content from Autor and Dorn (2013) is identical to "eye-hand-foot coordination". Task measures are collapsed from the occ1990dd level to the level of our modified consistent occupation codes, occ1990dd_18, which allow for extension to 2018. Panel A uses the inverse hyperbolic sine (IHS) of patent counts and the original ten-point scale for the task measures. Panel B uses year-specific percentiles of both patent counts and task content.

Table A9: Cross-Sectional Relationship between Occupational Augmentation and Automation Exposure and Occupational Task Content

	Routine (1)	Abstract (2)	Eye-hand-foot coordination (3)	Finger dexterity (4)	Direction, control and planning (5)	Set limits, and tolerances, and standards (6)	GED math (7)
				A. Occupational task content			
AugX	-0.206***	0.027	-0.007	-0.181***	0.136*	-0.230***	-0.082**
(Pat Count IHS, Occ-Link)	(0.044)	(0.041)	(0.020)	(0.030)	(0.058)	(0.063)	(0.028)
AutomX	0.556***	-0.256***	0.095***	0.292***	-0.431***	0.820***	-0.080*
(Pat Count IHS)	(0.037)	(0.055)	(0.021)	(0.025)	(0.082)	(0.056)	(0.033)
R^2	0.344	0.094	0.044	0.280	0.105	0.317	0.074
				B. Occupational task content percentile			
AugX	-0.369***	0.009	0.033	-0.407***	0.215**	-0.299***	-0.126*
(Pat Count Pctl, Occ-Link)	(0.074)	(0.064)	(0.075)	(0.069)	(0.069)	(0.073)	(0.055)
AutomX	0.841***	-0.341***	0.247***	0.708***	-0.560***	0.769***	-0.149**
(Pat Count Pctl)	(0.050)	(0.063)	(0.059)	(0.049)	(0.065)	(0.050)	(0.054)
R^2	0.377	0.109	0.067	0.267	0.190	0.358	0.071

Notes: Estimates pool patent counts from 1980 to 2018. All models are weighted by year-specific occupational employment share, include year fixed effects, and have 1,212 observations. Columns 1 and 2 use the routine and abstract task content measures defined in Autor and Dorn (2013) and columns 4-7 use the task measures defined in Autor et al. (2003). $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

D Constructing demand shifts

Here, we outline in more detail how we construct demand shifts relating to changes in import competition from China, and to population aging.

D.1 Constructing demand shift from Chinese import competition

We obtain import data for manufacturing industries classified by consistent SIC 87 codes ('SIC87dd') over 1991–2014 from Autor et al. (2013). We crosswalk these to consistent Census industries ('ind1990ddx') using 1991 value-added weights obtained from the NBER-CES Manufacturing Industry Database. Since employment data are observed by a different consistent Census industry, 'ind1990dd18' which is mostly but not entirely compatible with ind1990ddx codes, we create a new classification ('ind1990ddx18') which aggregates manufacturing categories where needed. This gives us a balanced panel of 69 manufacturing industries.

We retain years 1991, 2000, and 2014, and construct long differences over 1991–2000 and 2000–2014, scaling these up to match the time periods of our other data (1990–2000 and 2000–2018). For each industry, this gives us two changes in import competition, defined as changes in Chinese imports for other developed countries ($\Delta M_{i,t}^{OC}$) divided by the industry's U.S. market size in 1988 (U.S. industry output plus imports minus exports, $Y_{i,1988} + M_{i,1988} - E_{i,1988}$).

Using this, we then construct occupational exposure to these changes in import competition, as seen in equation 27. Note that for non-manufacturing industries, the values are set to zero, such that an occupation's exposure depends on exposure to manufacturing industries as well as exposure to manufacturing industries that have experienced different amounts of import competition from China. Our regression models control for occupational employment shares in manufacturing.

D.2 Constructing demand shift from population aging

We use Bureau of Labor Statistics' Consumption Expenditure (CE) data over 2002–2018 combined with Census population data over 1980–2018 to predict the annual demand for each Uniform Commercial Code (UCC) product category over 1970–2018 (largely following DellaVigna and Pollet (2007)), and then crosswalk these predictions to consistent industries (ind1990dd_18) to obtain predicted consumption by industry. The main text describes how these consumption levels by industry are then used to construct occupational exposure to industry demand shifts.

For each UCC category (k), we take the following steps.

1. **Annualize consumption.** The CE data are an unbalanced panel, because each consumption unit (CU)– effectively a household– does not occur in every monthly survey. We therefore use the twelve monthly CE surveys within each calendar year

and scale up the recorded consumption of the CU by 12/(number of months the CU appears in the survey).

2. **Pool data years.** We pool the annualized CE data across 2002–2018 because some UCC categories are not present in all years – we scale consumption for each UCC by the number of years they are observed. This yields c_{ik}, the average annual consumption for consumption unit i and UCC product category k.

3. **Estimate age-consumption profiles.** Next, we estimate the age-consumption profile relating consumption by consumption unit i and product category k to the household structure observed in the CE data:

$$c_{ik} = \sum_j \beta_{jk} H_{ij} + \sum_j \gamma_{jk} S_{ij} + \sum_j \delta_{jk} O_{ij} + \varepsilon_{ik}$$

where H_{ij} is the dummy indicating whether household i has a head in age bin j, S_{ij} is a dummy indicating whether household i has a spouse in age bin j, and O_{ij} is the number of other people (i.e. other than head or spouse) of household i in age bin j, and ε_{ik} is the error term. Note that this regression has no intercept, such that the coefficients can be interpreted as consumption per household member. We estimate this model separately for each UCC product category and weight models by population weights. Note that pooling data across years assumes consumption profiles by age are stable over time: this is supported by DellaVigna and Pollet (2007)'s analysis.

4. **Calculate household age shares.** We estimate year-averaged shares of head, spouse, and other people using population weights available in CE data:

$$h_j = \frac{\sum_i \text{Nr of heads in CU } i \text{ in age bin } j \times \text{CU } i\text{'s pop weight}}{\sum_i \text{Nr of total members of CU } i \text{ in age bin } j \times \text{CU } i\text{'s pop weight}}$$
$$s_j = \frac{\sum_i \text{Nr of spouses in CU } i \text{ in age bin } j \times \text{CU } i\text{'s pop weight}}{\sum_i \text{Nr of total members of CU } i \text{ in age bin } j \times \text{CU } i\text{'s pop weight}}$$
$$o_j = \frac{\sum_i \text{Nr of other people in CU } i \text{ in age bin } j \times \text{CU } i\text{'s pop weight}}{\sum_i \text{Nr of total members of CU } i \text{ in age bin } j \times \text{CU } i\text{'s pop weight}}$$

5. **Predict consumption.** Here we combine the estimated age-consumption coefficients and household share data (constructed above in steps 3 and 4, respectively) with Census population data over 1980–2018 to obtain aggregate predictions of consumption:

$$\widehat{c}_{kt} = \sum_j N_{jt} \times (\widehat{\beta}_{j,k} h_j + \widehat{\gamma}_{j,k} s_j + \widehat{\delta}_{jk} o_j),$$

where N_{jt} is the total U.S. population within the age bin j in year t. As such, $\widehat{c}_{kt}$ is predicted consumption for product category k in year t, based on the changing age distribution of the population over 1980–2018.

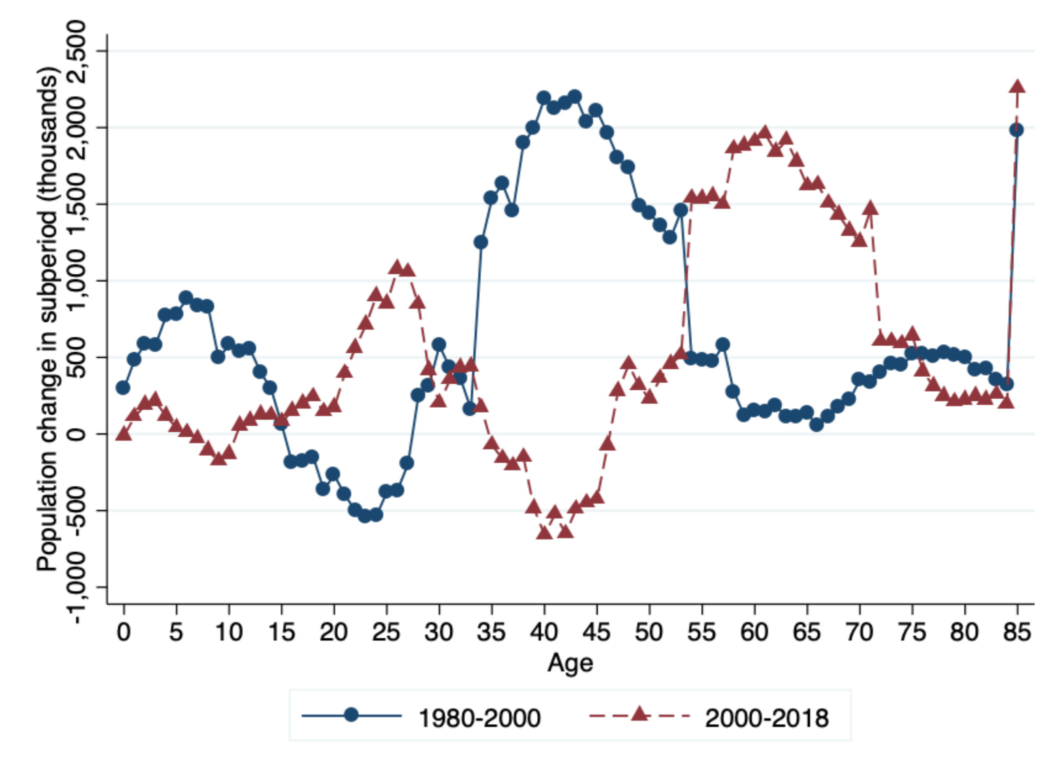
6. **Crosswalk predictions to consistent Census industries.** As a final step, we crosswalk these predictions to consistent Census industries. The crosswalk path is as follows: CE → PCE 2017 → BEA commodity 2012 → BEA industry 2012 → NAICS 2012 → NAICS 2007 → CIC 2010 → CIC 1990 → ind1990dd → ind1990dd_18, where

    - CE is the Consumer Expenditure Survey;
    - PCE 2017 are 2017 Personal Consumption Expenditures;
    - BEA commodity 2012 are 2012 Bureau of Economic Analysis commodity codes;
    - BEA industry 2012 are 2012 Bureau of Economic Analysis industry codes;
    - NAICS 2012 are 2012 North American Industry Classification System codes;
    - NAICS 2007 are 2007 North American Industry Classification System;
    - CIC 2010 are 2010 Census industry codes;
    - CIC 1990 are 1990 Census industry codes;
    - ind1990dd are consistent industry codes constructed by David Dorn; and
    - ind1990dd_18 are our modified version of these codes to allow extension of the panel to 2018.

    To link CE to PCE we use the weights indicated by the BLS. PCE categories that match to multiple BEA commodities are split using weights generated by producer value. Using producer values allows us to manually include the trade and transportation margins from the BEA use table when crosswalking PCE to BEA commodity codes. This adjustment allows us to avoid dropping retail and wholesale commodities. In all other crosswalks, expenditures are split evenly when one category matches to multiple categories. In our baseline demand shift results shown in Table 5, we used full input-output adjustments since industry demands intrinsically have an input-output component. Our results are robust to using demand shifts without input-output linkages, i.e. equating BEA commodity and industry codes.

Figure A2 shows changes in the population by age over 1980–2000 and 2000–2018, highlighting the importance of the aging Baby Boom generation. In the first period, this cohort was prime-aged and having children, also leading to an increase for ages 0 to 10. Over the subsequent two decades, this cohort entered older age groups, creating a large spike at ages 55 and above, as well as a smaller increase in the number of young adults.

Figure A2: U.S. Population Change by Age, 1980–2000 and 2000–2018



E Robustness checks

This Appendix presents several robustness checks on our results: these are referenced and described in the main text.

Table A10: Occupational Employment in New Work and Augmentation Exposure, 1940–2018

	(1)	(2)	(3)	(4)	(5)
	A. 100 × IHS Employment in New Work, Occupation-Industry Cells				
AugX	6.83***	7.74***	4.49*	9.43***	9.95***
(Pat Count IHS, Occ-Link)	(2.02)	(1.86)	(2.12)	(2.06)	(1.97)
N	328,715	328,715	328,715	328,715	328,715
R^2	0.475	0.572	0.495	0.537	0.623
	B. Percentile of Employment in New Work, Occupation-Industry Cells				
AugX	0.141***	0.153***	0.150***	0.182***	0.192***
(Pat Count Pctl, Occ-Link)	(0.031)	(0.031)	(0.032)	(0.031)	(0.031)
N	328,715	328,715	328,715	328,715	328,715
R^2	0.412	0.471	0.443	0.461	0.516
Broad Ind × Year FE	X			X	
Ind × Year FE		X			X
Broad Occ × Year FE			X	X	X

Notes: Census occupations and industries over 1940–2018. Models weighted by annual occupation-industry employment shares. Broad industries and occupations are each 12 groups consistently defined over time. Standard errors clustered by occupation-year in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table A11: Occupational Employment Growth and Augmentation versus Automation Exposure, 1980–2018

Dependent variable: 100 × DHS Employment Change in Occupation-Industry Cells

	(1)	(2)	(3)	(4)	(5)	(6)
	A. 1980–2018 Long Difference					
AugX	0.26**	0.44***			0.51***	0.50***
(Pat Count Pctl, Ind×Occ-Link)	(0.08)	(0.08)			(0.08)	(0.08)
AutomX			-0.61***	-0.25*	-0.71***	-0.33**
(Pat Count Pctl)			(0.07)	(0.10)	(0.07)	(0.10)
N	42,055	42,055	42,055	42,055	42,055	42,055
R^2	0.43	0.49	0.45	0.48	0.45	0.49
	B. 1980–2000 & 2000–2018 Stacked First Differences					
AugX	0.15**	0.29***			0.32***	0.33***
(Pat Count Pctl, Ind×Occ-Link)	(0.06)	(0.05)			(0.05)	(0.05)
AutomX			-0.35***	-0.13*	-0.41***	-0.19**
(Pat Count Pctl)			(0.04)	(0.06)	(0.04)	(0.06)
N	81,328	81,328	81,328	81,328	81,328	81,328
R^2	0.34	0.37	0.35	0.37	0.35	0.37
Broad Occ (× Year) FE		X		X		X
Ind (× Year) FE	X	X	X	X	X	X

Notes: Consistently defined Census occupations and industries over 1980–2018. Models weighted by average annual occupation-industry cell employment shares at the start and end of the time period. Broad occupations and industries are 12 and 13 groups consistently defined over time. Standard errors clustered by occupation-by-industry in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table A12: Occupational Wagebill Growth and Augmentation versus Automation Exposure, 1980–2018

Dependent variable: 100 × DHS Employment/Wagebill Change in Occupation-Industry Cells

	Δ Employment		Δ Wagebill		Δ E[Wagebill] - Δ Employment		Δ Adj. Wagebill - Δ Employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. 1980–2018 Long Difference							
AugX	0.51***	0.50***	0.43***	0.44***	-0.09***	-0.07***	0.01	0.01
(Pat Count Pctl, Ind×Occ-Link)	(0.08)	(0.08)	(0.08)	(0.08)	(0.01)	(0.01)	(0.01)	(0.01)
AutomX	-0.71***	-0.33**	-0.75***	-0.27**	-0.01	0.05***	-0.03*	0.01
(Pat Count Pctl)	(0.07)	(0.10)	(0.07)	(0.10)	(0.01)	(0.01)	(0.01)	(0.02)
N	42,055	42,055	42,055	42,055	42,055	42,055	42,055	42,055
R^2	0.45	0.49	0.47	0.52	0.40	0.48	0.40	0.45
	B. 1980–2000 & 2000–2018 Stacked First Differences							
AugX	0.32***	0.33***	0.26***	0.29***	-0.05***	-0.04***	-0.00	0.00
(Pat Count Pctl, Ind×Occ-Link)	(0.05)	(0.05)	(0.05)	(0.05)	(0.01)	(0.01)	(0.01)	(0.01)
AutomX	-0.41***	-0.19**	-0.43***	-0.15*	-0.01	0.03***	-0.02*	0.00
(Pat Count Pctl)	(0.04)	(0.06)	(0.05)	(0.06)	(0.01)	(0.01)	(0.01)	(0.01)
N	81,328	81,328	81,328	81,328	81,328	81,328	81,328	81,328
R^2	0.35	0.37	0.37	0.40	0.44	0.50	0.33	0.37
Broad Occ (× Year) FE		X		X		X		X
Ind (× Year) FE	X	X	X	X	X	X	X	X

Notes: Dependent variable is Davis-Haltiwanger-Schuh (DHS) change in employment (columns 1-2), DHS change in wagebill (columns 3-4), expected DHS change in wagebill net of employment change (columns 5-6), expected DHS change in composition-adjusted wagebill net of employment change (columns 7-8). Consistently defined Census occupations and industries over 1980–2018. Broad occupations and industries are 12 and 13 groups consistently defined over time. Standard errors clustered by occupation-by-industry in parentheses. $^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

F Proofs to propositions

Proof to Proposition 1

The price index P_j is the minimum cost for buying an additional unit of good j in equilibrium. Given that equation (4) is CES, the corresponding expression for the marginal cost of producing Y_j is given by:

$$P_j = \left[\int_{N_j-1}^{N_j} p_j(i)^{1-\sigma} di\right]^{\frac{1}{1-\sigma}} \tag{A1}$$

Using equation (8), equation (A1) can be written as:

$$\begin{aligned} P_j^{1-\sigma} &= \int_{N_j-1}^{N_j} R_j^{[1-\eta][1-\sigma]} di + \int_{I_j}^{N_j} \left[\frac{W_j}{\gamma_j(i)}\right]^{[1-\eta][1-\sigma]} di \\ &= [I_j - N_j + 1] R_j^{1-\hat{\sigma}} + W_j^{1-\hat{\sigma}} \int_{I_j}^{N_j} \gamma_j(i)^{\hat{\sigma}-1} di \end{aligned} \tag{A2}$$

with $[1-\eta][1-\sigma] = 1-\hat{\sigma}$ given that $\hat{\sigma} \equiv [1-\eta]\sigma + \eta$.

We can rewrite the factor-market clearing conditions (equations (12) and (13)) to solve for W_j and R_j in equilibrium:

$$W_j = \left[\frac{[1-\eta]\beta_j Y P_j^{\sigma-1}}{L_j} \int_{I_j}^{N_j} \gamma_j(i)^{\hat{\sigma}-1} di\right]^{\frac{1}{\hat{\sigma}}} \tag{A3}$$

and

$$R_j = \left[\frac{[1-\eta]\beta_j Y P_j^{\sigma-1}}{K_j} [I_j - N_j + 1]\right]^{\frac{1}{\hat{\sigma}}} \tag{A4}$$

Substituting the expressions for W_j and R_j from equations (A3) and (A4) into equation (A2) gives:

$$\begin{aligned} P_j^{1-\sigma} =& [I_j - N_j + 1] \left[\frac{[1-\eta]\beta_j Y P_j^{\sigma-1}}{K_j} [I_j - N_j + 1]\right]^{\frac{1-\hat{\sigma}}{\hat{\sigma}}} \\ &+ \left[\frac{[1-\eta]\beta_j Y P_j^{\sigma-1}}{L_j} \int_{I_j}^{N_j} \gamma_j(i)^{\hat{\sigma}-1} di\right]^{\frac{1-\hat{\sigma}}{\hat{\sigma}}} \int_{I_j}^{N_j} \gamma_j(i)^{\hat{\sigma}-1} di \end{aligned} \tag{A5}$$

Bringing the P_j on the left-hand side to the right-hand side and using that $1/[1-\eta]-1 = \eta/[1-\eta]$ gives:

$$[1-\eta] Y_j = P_j^{\frac{\eta}{1-\eta}} \left[[I_j - N_j + 1]^{\frac{1}{\hat{\sigma}}} K_j^{\frac{\hat{\sigma}-1}{\hat{\sigma}}} + [\int_{I_j}^{N_j} \gamma_j(i)^{\hat{\sigma}-1} di]^{\frac{1}{\hat{\sigma}}} L_j^{\frac{\hat{\sigma}-1}{\hat{\sigma}}}\right]^{\frac{\hat{\sigma}}{\hat{\sigma}-1}} \tag{A6}$$

Proof of Proposition 2

The wagebills must satisfy:

$$W_L L_U = \alpha^U s_U^L P_U Y_U = \alpha^U s_U^L \beta Y \tag{A7}$$
$$W_L L_S = \alpha^S s_S^L P_S Y_S = \alpha^S s_S^L [1-\beta] Y \tag{A8}$$
$$W_H H_U = (1-\alpha^U) s_U^L P_U Y_U = (1-\alpha^U) s_U^L \beta Y \tag{A9}$$
$$W_H H_S = (1-\alpha^S) s_S^L P_S Y_S = (1-\alpha^S) s_S^L [1-\beta] Y \tag{A10}$$

where W_L and W_H are the respective wages for low-skilled and high-skilled workers. Acemoglu and Restrepo (2019) show that the labor share in sector j is given by:

$$s_j^L = \left[1 + \left[\frac{1-\Gamma_j}{\Gamma_j}\right]^{\frac{1}{\sigma}} \left[\frac{K_j}{\mathcal{L}_j}\right]^{\frac{\sigma-1}{\sigma}}\right]^{-1} \tag{A11}$$

with σ the elasticity between tasks in goods production and

$$\Gamma_j \equiv \frac{\int_{I_j}^{N_j} \gamma^L(i)^{\sigma-1} di}{[I - N_j + 1]^{\sigma-1} + \int_{I_j}^{N_j} \gamma_j(i)^{\sigma-1} di}. \tag{A12}$$

In the context of our model, $\mathcal{L}_j$ denotes the labor composite $\mathcal{L}_j = (L_j)^{\alpha_j}(H_j)^{1-\alpha_j}$. Differentiating the expression for the labor share,

$$d\ln(s_j^L) = \frac{1}{\sigma}\frac{1-s_j^L}{1-\Gamma_j} d\ln(\Gamma_j) - \frac{\sigma-1}{\sigma}\left[1-s_j^L\right] d\ln\left(\frac{K_j}{\mathcal{L}_j}\right) \tag{A13}$$

with $d\ln(\Gamma_j) < 0$ for automation (i.e. $I_j > 0$); $d\ln(\Gamma_j) > 0$ for augmentation leading to the creation of new labor-intensive tasks (i.e. $N_j > 0$); and $\mathcal{L}_j$ is defined as above. Note that $d\ln K_j = 0$ in our model, as capital is sector-specific and its supply is inelastic.

Differentiating (A7), we obtain:

$$\begin{aligned} d\ln(W_L) + d\ln(L_U) = &\frac{1}{\sigma}\frac{1-s_U^L}{1-\Gamma_U} d\ln(\Gamma_U) \\ &+ \frac{\sigma-1}{\sigma}[1-s_U^L]\left[\alpha^U d\ln(L_U) + (1-\alpha^U) d\ln(H_U)\right] \\ &+ d\ln(Y) \end{aligned} \tag{A14}$$

with $d\ln(\Gamma_U) < 0$. Note that the term $\left[\alpha^U d\ln(L_U) + (1-\alpha^U) d\ln(H_U)\right]$ is obtained from the Cobb-Douglas labor aggregate in sector U, $L_U = (L_U)^{\alpha^U}(H_U)^{1-\alpha^U}$.

For low-skilled workers in sector S, equation (A8), where no automation is occurring (i.e.,

$d\ln(\Gamma_S)=0$):

$$d\ln(W_L)+d\ln(L_S)=\frac{\sigma-1}{\sigma}[1-s_S^L]\left[\alpha^S d\ln(L^S)+(1-\alpha^S)d\ln(H_S)\right] + d\ln(Y) \tag{A15}$$

Equating low-skilled worker wage changes across the two sectors and denoting $T\equiv\frac{1}{\sigma}\frac{1-s_{LU}}{1-\Gamma_U}d\ln(\Gamma_U)$, as before,

$$\left[1-\frac{\sigma-1}{\sigma}[1-s_U^L]\alpha^U\right]d\ln(L_U)-\frac{\sigma-1}{\sigma}[1-s_U^L]\left[(1-\alpha^U)d\ln(H_U)\right]-T=$$
$$\left[1-\frac{\sigma-1}{\sigma}[1-s_S^L]\alpha^S\right]d\ln(L_S)-\frac{\sigma-1}{\sigma}[1-s_S^L]\left[(1-\alpha^S)d\ln(H_S)\right]$$

We repeat this for high-skilled workers, given by (A9) and (A10). Consolidating our equations, we have:

$$\left[1-\frac{\sigma-1}{\sigma}(1-s_U^L)(1-\alpha^U)\right]d\ln H_U-\alpha^U\frac{\sigma-1}{\sigma}(1-s_U^L)d\ln L_U-T$$
$$=\left[1-\frac{\sigma-1}{\sigma}(1-s_S^L)(1-\alpha^S)\right]d\ln H_S-\alpha^S\frac{\sigma-1}{\sigma}(1-s_S^L)d\ln L_S \tag{A16}$$

and

$$\left[1-\alpha^U\frac{\sigma-1}{\sigma}(1-s_U^L)\right]d\ln L_U-\frac{\sigma-1}{\sigma}(1-s_U^L)(1-\alpha^U)d\ln H_U-T$$
$$=\left[1-\alpha^S\frac{\sigma-1}{\sigma}(1-s_S^L)\right]d\ln L_S-\frac{\sigma-1}{\sigma}(1-s_S^L)(1-\alpha^S)d\ln H_S \tag{A17}$$

with $dL_U=-dL_S$ and $dH_U=-dH_S$. Noting the need to change $d\ln x$ to dx/x, and solving the system of four equations with four unknowns, we arrive at expressions for dH_U and dL_U:

$$dH_U=\frac{TL}{\frac{1}{\sigma}\left[\frac{L}{H_U}+\frac{L}{H-H_U}\right]+\frac{\sigma-1}{\sigma}\left[s_U^L\frac{L}{H_U}+s_S^L\frac{L}{H-H_U}+\left(\frac{L_U}{H_U}-\frac{L-L_U}{H-H_U}\right)(\alpha^U(1-s_U^L)-\alpha^S(1-s_S^L))\right]} \tag{A18}$$

$$dL_U=\frac{TH}{\frac{1}{\sigma}\left(\frac{H_SL}{L_UL_S}+\frac{H_UL}{L_UL_S}\right)+\frac{\sigma-1}{\sigma}\left[s_U^L\frac{H_SL}{L_UL_S}+s_S^L\frac{H_UL}{L_UL_S}+\left(\frac{H_S}{L_S}-\frac{H_U}{L_U}\right)(\alpha^U(1-s_U^L)-\alpha^S(1-s_S^L))\right]} \tag{A19}$$

where the relationship between dH_U and dL_U is given by:

$$dL_U=\frac{\frac{1}{H_U}+\frac{1}{H_S}}{\frac{1}{L_U}+\frac{1}{L_S}}dH_U$$

Next, we will characterize the sign of the above equations. First, focus on equation (A18). As $T < 0$, the numerator is negative. The first term of the denominator is positive, since labor quantities and the elasticity of substitution σ are non-negative. The second term of denominator varies linearly in s_U^L and s_S^L. Its derivative with respect to s_U^L and s_S^L are given by:

$$\frac{L_U}{H_U} + \frac{L_S}{H_U} - \alpha^U \left(\frac{L_U}{H_U} - \frac{L - L_U}{H - H_U} \right) > 0$$
$$\frac{L}{H - H_U} + \alpha^S \left(\frac{L_U}{H_U} - \frac{L - L_U}{H - H_U} \right) > 0$$

We can then look at the cases of $s_U^L, s_S^L \to 0$ and $s_U^L, s_S^L \to 1$. In the first case, the second term of the denominator collapses to

$$\frac{\sigma - 1}{\sigma} \left(\frac{L_U}{H_U} - \frac{L - L_U}{H - H_U} \right) (\alpha^U - \alpha^S) > 0$$

where the inequality is a consequence of two facts: 1) the ratio of unskilled to skilled workers in sector U being greater than that ratio in sector S; and 2) $\alpha^U - \alpha^S > 0$. In the second case, the second term of the denominator collapses to

$$\frac{\sigma - 1}{\sigma} \left[\frac{L}{H_U} + \frac{L}{H - H_U} \right] > 0$$

Using the fact that the second term of the denominator varies linearly in s_U^L and s_S^L, the signs of the two cases implies that the second term of the denominator is positive. Together, as the denominator is a convex combination of two positive quantities, it too is positive, implying that $dH_U < 0$. Applying (F) yields $dL_U < 0$.

To prove the final part of the proposition – that the derivatives have the opposite sign when augmentation or automation occurs in sector S instead of sector U – we differentiate (A7)-(A10):

$$d \ln W_L + d \ln L_U = \frac{\sigma - 1}{\sigma}(1 - s_U^L) d \ln L_U + d \ln Y \tag{A20}$$
$$d \ln W_H + d \ln L_S = \frac{1}{\sigma} \frac{1 - s_S^L}{1 - \Gamma_S} d \ln \Gamma_S + \frac{\sigma - 1}{\sigma}(1 - s_U^L) d \ln L_U + d \ln Y \tag{A21}$$
$$d \ln W_L + d \ln H_U = \frac{\sigma - 1}{\sigma}(1 - s_U^L) d \ln L_U + d \ln Y \tag{A22}$$
$$d \ln W_H + d \ln H_S = \frac{1}{\sigma} \frac{1 - s_S^L}{1 - \Gamma_S} d \ln \Gamma_S + \frac{\sigma - 1}{\sigma}(1 - s_U^L) d \ln L_U + d \ln Y \tag{A23}$$

Noting the differences between (A14)-(A15) and (A20)-(A21) are the switching of subscripts U and S on the right-hand side, which also extends to (A22) and (A23), there is a symmetry between sectors U and S, conveniently allowing us to transfer the results above to the case in which automation and augmentation occur in sector S.

Proof of Proposition 3

As a first step, note that Assumption 3 implies

$$\frac{W_j}{\gamma(I^*)} > R > \frac{W_j}{\gamma(I(N))}$$

so that it is always profitable to automate and create new tasks. In the initial equilibrium, we will have that $V_j^N = V_j^I$ (see Lemma 1). Differentiating the value functions of sector j with respect to β_j, we obtain the effect of a positive demand shift on incentives for automation and new task creation in sector j:

$$\frac{\partial V_j^I}{\partial \beta_j} = \underbrace{\frac{\partial Y_j P_j^\sigma}{\partial \beta_j}}_{A>0} \times \underbrace{(1-\mu)\eta\left[R_j^{1-\hat{\sigma}} - \left(\frac{W_j}{\gamma_j(I)}\right)^{1-\hat{\sigma}}\right]}_{B>0} + \underbrace{\frac{\partial\left[R_j^{1-\hat{\sigma}} - \left(\frac{W_j}{\gamma_j(I)}\right)^{1-\hat{\sigma}}\right]}{\partial \beta_j}}_{C<0} \times \underbrace{(1-\mu)\eta Y_j P_j^\sigma}_{D>0} \tag{A24}$$

$$\frac{\partial V_j^N}{\partial \beta_j} = \underbrace{\frac{\partial Y_j P_j^\sigma}{\partial \beta_j}}_{A>0} \times \underbrace{(1-\mu)\eta\left[\left(\frac{W_j}{\gamma_j(N)}\right)^{1-\hat{\sigma}} - R_j^{1-\hat{\sigma}}\right]}_{E>0} + \underbrace{\frac{\partial\left[\left(\frac{W_j}{\gamma_j(N)}\right)^{1-\hat{\sigma}} - R_j^{1-\hat{\sigma}}\right]}{\partial \beta_j}}_{F>0} \times \underbrace{(1-\mu)\eta Y_j P_j^\sigma}_{D>0} \tag{A25}$$

In both equations, A is positive as output and prices increase from the outward demand shift. For the value of additional task automation, A multiplies a term (B) which is positive by Assumption 3, as an increase in the range of tasks that are automated increases productivity. Similarly, for the value of additional augmentation, A multiplies a term (E) which is positive by Assumption 2, as an increase in new tasks also increases productivity. Since rental rates rise more strongly than wages do, C is negative, and F is positive; and both C and F multiply a positive term (D). Therefore, the incentives for new task creation in the sector with the demand expansion are unambiguously positive and exceed incentives for task automation in this sector, $\frac{\partial V_j^N}{\partial \beta_j} > \frac{\partial V_j^I}{\partial \beta_j}$, if in the initial equilibrium $V_j^N = V_j^I$ since this implies $B = E$.

We can also study incentives for automation and new task creation in the other sector, denoted by $\tilde{\jmath} \neq j$, when β_j increases. This illuminates the effects of a demand contraction in sector $\tilde{\jmath}$. Differentiating,

$$\frac{\partial V_{\tilde{\jmath}}^{I}}{\partial \beta_j} = \underbrace{\frac{\partial Y_{\tilde{\jmath}} P_{\tilde{\jmath}}^{\sigma}}{\partial \beta_j}}_{G<0} \times \underbrace{(1-\mu)\eta \left[R_{\tilde{\jmath}}^{1-\hat{\sigma}} - \left(\frac{W_{\tilde{\jmath}}}{\gamma_{\tilde{\jmath}(I)}}\right)^{1-\hat{\sigma}}\right]}_{H>0} + \underbrace{\frac{\partial \left[R_{\tilde{\jmath}}^{1-\hat{\sigma}} - \left(\frac{W_{\tilde{\jmath}}}{\gamma_{\tilde{\jmath}}(I)}\right)^{1-\hat{\sigma}}\right]}{\partial \beta_j}}_{I>0} \times \underbrace{(1-\mu)\eta Y_{\tilde{\jmath}} P_{\tilde{\jmath}}^{\sigma}}_{J>0} \tag{A26}$$

$$\frac{\partial V_{\tilde{\jmath}}^{N}}{\partial \beta_j} = \underbrace{\frac{\partial Y_{\tilde{\jmath}} P_{\tilde{\jmath}}^{\sigma}}{\partial \beta_j}}_{G<0} \times \underbrace{(1-\mu)\eta \left[\left(\frac{W_{\tilde{\jmath}}}{\gamma_{\tilde{\jmath}}(N)}\right)^{1-\hat{\sigma}} - R_{\tilde{\jmath}}^{1-\hat{\sigma}}\right]}_{K>0} + \underbrace{\frac{\partial \left[\left(\frac{W_{\tilde{\jmath}}}{\gamma_{\tilde{\jmath}}(N)}\right)^{1-\hat{\sigma}} - R_{\tilde{\jmath}}^{1-\hat{\sigma}}\right]}{\partial \beta_j}}_{L<0} \times \underbrace{(1-\mu)\eta Y_{\tilde{\jmath}} P_{\tilde{\jmath}}^{\sigma}}_{J>0} \tag{A27}$$

Hence in response to a demand contraction, incentives for new task creation in the sector are reduced: both overall, $\frac{\partial V_{\tilde{\jmath}}^{N}}{\partial \beta_j} < 0$, and relative to automation in the sector $\frac{\partial V_{\tilde{\jmath}}^{N}}{\partial \beta_j} < \frac{\partial V_{\tilde{\jmath}}^{I}}{\partial \beta_j}$. We can summarize the relative magnitudes of the changes in the value of innovations in response to changes in demand as follows:

$$\frac{\partial V_{j}^{N}}{\partial \beta_j} > \frac{\partial V_{j}^{I}}{\partial \beta_j}$$
$$\frac{\partial V_{\tilde{\jmath}}^{N}}{\partial \beta_j} < \frac{\partial V_{\tilde{\jmath}}^{I}}{\partial \beta_j}$$

Because entrepreneurs' wages in a given sector-innovation cell are equal to the value of the innovations they create intermediates for ($w_j^m = V_j^m$), and since $\Delta I^j = E_I^j$ and $\Delta N^j = E_N^j$, we obtain that

$$\frac{\partial \Delta N_j}{\partial \beta_j} > \frac{\partial \Delta I_j}{\partial \beta_j}$$
$$\frac{\partial \Delta N_{\tilde{\jmath}}}{\partial \beta_j} < \frac{\partial \Delta I_{\tilde{\jmath}}}{\partial \beta_j}$$

which corresponds to our proposition.