4-2020

The Evolution of Work in the United States

Enghin Atalay

Phai Phongthiengtham

Sebastian Sotelo

Daniel Tannenbaum

American Economic Journal: Applied Economics 2020, 12(2): 1–34
https://doi.org/10.1257/app.20190070

The Evolution of Work in the United States†

By Enghin Atalay, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum*

Using the text from job ads, we introduce a new dataset to describe the evolution of work from 1950 to 2000. We show that the transformation of the US labor market away from routine cognitive and manual tasks and toward nonroutine interactive and analytic tasks has been larger than prior research has found, with a substantial fraction of total changes occurring within narrowly defined job titles. We provide narrative and systematic evidence on changes in task content within job titles and on the emergence and disappearance of individual job titles. (*JEL* E24, J21, J24, J31, N32)

The dramatic technological innovations of the twentieth century and the rise of international offshoring have transformed labor markets (Autor 2015; Brynjolfsson and McAfee 2014; Firpo, Fortin, and Lemieux 2014). The ensuing decline in the real earnings of low-skilled workers, the widening earnings distribution, and the hollowing out of middle-skilled jobs have turned the attention of policymakers and researchers to a detailed study of the activities that workers do on the job. Despite substantial recent progress, however, measuring changes in available jobs and their associated tasks remains a challenge.

One approach to measuring the changing nature of work that is widely used in the literature is to study the occupational shares of US employment. Using this approach, the literature has identified a dramatic transformation in the US labor market. For example, occupations that are intensive in routine tasks have shrunk as a share of total employment, while those that emphasize nonroutine tasks as well as social and cognitive skills have grown (Autor, Levy, and Murnane

*Atalay: Federal Reserve Bank of Philadelphia, 10 N. Independence Mall West, Philadelphia, PA 19106 (email: atalayecon@gmail.com); Phongthiengtham: IBM Corporation, 1 North Castle Drive, Armonk, NY 10504 (email: phai@ibm.com); Sotelo: Department of Economics, University of Michigan-Ann Arbor, 611 Tappan Avenue, Ann Arbor, MI 48109 (email: ssotelo@umich.edu); Tannenbaum: Department of Economics, University of Nebraska-Lincoln, 730 N. 14th Street, Lincoln, NE 68588 (email: dtannenbaum@unl.edu). David Deming was coeditor for this article. We thank Sara Heller, Andrei Levchenko, Pablo Ottonello, Pascual Restrepo, and Chris Taber for constructive and helpful comments, and Barbara Nordin and Erin Robertson for excellent editorial assistance. We also thank seminar audiences at APPAM, ASSA, BYU, the University of British Columbia, Cornell, the Federal Reserve Bank of Richmond, University of Geneva, Iowa State University, University of Michigan, Penn State, PEA Congress, the NBER Summer Institute, the Upjohn Institute, Virginia Tech, and the University of Wisconsin. We acknowledge financial support from the Washington Center for Equitable Growth, the Upjohn Institute, and Grant #92-18-05 from the Russell Sage Foundation. This paper was previously circulated under the title "The Evolving U.S. Occupational Structure." Research results and conclusions expressed are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve System, or the Federal Reserve Board of Governors.
†Go to https://doi.org/10.1257/app.20190070 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

2003; Autor, Katz, and Kearney 2006; Autor and Dorn 2013; Deming 2017).[1] The evidence is largely silent, however, on whether the occupations themselves have changed. Data sources that are widely used to study the labor market in the United States are not well suited for studying task changes within occupations over time. Meanwhile, case studies—including those of managers (National Research Council 1999), production workers (Bartel, Ichniowski, and Shaw 2007), and cashiers (Basker, Klimek, and Van 2012)—document substantial transformations within individual occupations. These findings raise the question of whether these occupations are unique in experiencing changes in tasks or whether comparable changes have occurred elsewhere in the labor market.

In this paper, we introduce a new data source to document the transformation of job tasks in the United States. We construct our dataset from the text content of approximately 7.8 million job ads appearing in 3 major metropolitan newspapers: the *Boston Globe*, the *New York Times*, and the *Wall Street Journal*. We then map the words contained in the job descriptions to tasks. Our main strategy uses a mapping of words into routine and nonroutine tasks introduced by Spitz-Oener (2006), but for robustness, we consider alternative mappings, including one discussed by Deming and Kahn (2018). Since job ads that appear in newspapers do not contain Standard Occupational Classification (SOC) codes, we use machine-learning methods to map job titles (which we observe directly in the ads) to their corresponding SOCs.

To demonstrate that our job descriptions contain valuable information, we validate our new dataset in several ways. We show that our cross-sectional measures of occupations' task and skill measures correlate with those in the widely used Dictionary of Occupational Titles (DOT) and Occupational Information Network (O*NET). We further show that our new data are able to replicate, with broad concurrence, the key between-occupation trends in tasks documented by Autor, Levy, and Murnane (2003). We perform several checks on the data to provide evidence that neither the selection of ads into newspapers nor the fact that our data originate in metropolitan areas biases our results.[2]

Next, we show that substantial changes in job tasks have occurred since 1950. Using our database of newspaper ads, we demonstrate that words related to nonroutine tasks have been increasing in frequency, while words related to routine tasks (especially routine manual tasks) have declined in frequency between 1950 and 2000. The frequency of words related to routine cognitive tasks has declined by more than one-half over the sample period, from 2.0 mentions per 1,000 job ad words to 0.9 mentions per 1,000 words. The frequency of routine manual tasks has declined even more starkly. The frequency of words related to nonroutine analytic tasks, on the other hand, has increased from 2.9 to 5.5 mentions per 1,000 job ad

[1] Acemoglu and Autor (2011), among others, emphasize that skills and tasks refer to different work concepts: "A *task* is a unit of work activity that produces output (goods and services). In contrast, a skill is a worker's endowment of capabilities for performing various tasks" (Acemoglu and Autor 2011, 1045, emphasis in the original). We adopt these definitions of skills and tasks throughout our paper.

[2] First, we document that there are no trends over time in ad length or in the number of words that do not appear in the dictionary (online Appendix C.1). Second, we show that trends in the propensity of unemployed workers to search for jobs through help wanted ads do not vary with the task content of their prior occupation (online Appendix C.2). And third, we find evidence against these results being driven by the fact that most of our ads come from a selected number of large metro areas (online Appendix C.3).

words. Mentions of nonroutine interactive tasks have increased from 5.0 to 7.1 mentions per 1,000 job ad words. Since tasks have no natural unit of measurement, we consider an alternative measure of task changes that decomposes each job into its composite task shares. Both approaches lend support to these overall trends.

Our main finding is that a large share of the aggregate change in both nonroutine and routine tasks over our sample period has occurred *within* occupations rather than through changes in occupations' employment shares. In our benchmark decompositions, 88 percent of the overall changes in task content have occurred within rather than between job titles. We emphasize that the predominance of the within-occupation margin holds regardless of how finely one defines an occupation: four-digit SOC codes, six-digit SOC codes, or job titles. This finding is robust to alternative mappings between words and tasks and to alternative weighting methods and normalizations. Our finding is important because it implies that the transformation of the US labor market has been far more dramatic than previous research has found. It also suggests that fixing the task content of jobs at a point in time misses important features of the evolving nature of work in the United States and that standard data sources are unable to fully characterize this evolution.

We next provide new descriptive evidence on the evolution of individual job titles, a level of granularity unavailable in standard data sources. We confirm the finding of the National Research Council (1999) that managerial jobs in the United States have become much more interactive, emphasizing team building, coaching, and interactions with customers. We find similar changes for machinists and cashiers, along the lines of Bartel, Ichniowski, and Shaw (2007) and Basker, Klimek, and Van (2012). Next, we document the rise and fall of individual job titles. We find substantial turnover in the mix of job titles within six-digit SOCs. Further, we show that newer vintage job titles mention nonroutine analytic and interactive tasks more frequently and routine tasks less frequently. Taken together, these findings illustrate the many margins of change within occupations in the SOC classification—margins we could not observe before.

Our paper builds on two literatures. The first examines the causes and consequences of the evolution of occupations. Autor, Levy, and Murnane (2003) and Acemoglu and Autor (2011) develop the hypothesis that technological advances have reduced the demand for routine tasks, which in turn has led to a reduction in the wages of low- and middle-skilled workers. Deming (2017) documents that employment and wage growth has been confined to occupations that are intensive in both social and cognitive skills. Michaels, Rauch, and Redding (2018) study changes in employment shares by task content over a longer time horizon. They adopt a methodology related to ours, using verbs from the DOT's occupational descriptions and their thesaurus-based meanings. With the exception of Autor, Levy, and Murnane (2003), none of this prior work directly measures within-occupation variation across different editions of the DOT.[3]

[3] Autor, Levy, and Murnane (2003) use the 1977 and 1991 versions of the DOT to compare changes in occupations' task content and computer adoption rates. As they and Autor (2013) note, the update of the DOT was not exhaustive across occupations, potentially leading to status quo bias (Miller et al. 1980). We contrast occupational change measured in our data to what is possible using the DOT in online Appendix B.3 and also conclude that the DOT's ability to measure time-varying occupational tasks is limited.

Relative to this first literature, our paper contributes with a new measurement of time-varying characteristics of US occupations over the second half of the twentieth century. We introduce a new, publicly available dataset at the occupation-year level. This dataset includes measures of tasks, skill requirements, and other job characteristics between 1950 and 2000. Because they are built from newspaper text, our data rely on a continuously updated source and have the advantage over survey-based data of being collected in the field. Firms post these ads while they are actively searching for workers. We view this new dataset as complementary to data sources currently used to study the evolution of the US labor market. In particular, we extend what can be accomplished by linking across editions of the DOT or O*NET (Ross 2017). We also make our data available at the job title-year level, which measures tasks at a finer unit of analysis than the occupational level available in other datasets.[4]

Outside the US context, one paper that focuses on within-occupation changes is Spitz-Oener (2006), which uses survey data from four waves of German workers to track task changes within and between occupations from the late 1970s to the late 1990s. A comparable analysis in the United States cannot be achieved with existing surveys, and hence one of the contributions of this paper is to undertake the construction and validation of a new dataset that allows for such an analysis. Our newly constructed dataset covers a substantially longer period than the dataset used by Spitz-Oener (2006) and includes a much wider set of task and skill measures. The key takeaway from our analysis, resulting from our new measurement and framework, is that the evolution of job tasks in the United States has been even more dramatic than previously thought.[5]

The second literature on which our paper builds uses the text from online help wanted ads to study the labor market: how firms and workers match with one another, how firms differ in their job requirements, and how skill requirements have changed since the beginning of the Great Recession.[6] Using data from CareerBuilder, Marinescu and Wolthoff (forthcoming) document substantial variation in job ads' skill requirements and stated salaries within narrowly defined occupation codes. Also using online job ads, Hershbein and Kahn (2018) and Modestino, Shoag, and Ballance (2019) argue that jobs' skill requirements have increased during the post-Great Recession period; Deming and Kahn (2018) find that firms that post ads with a high frequency of words related to social and cognitive skills have higher labor productivity and pay higher wages.[7] Our contribution relative to this second

[4] Our new dataset can be found at https://occupationdata.github.io. Even though our job measures extend back to 1940, many of the exercises in this paper rely on mapping occupation codes across vintages of the decennial census, which is difficult to do for earlier periods. Our dataset has recently been applied by Anastasopoulos et al. (2018); Cortes, Jaimovich, and Siu (2018); and Deming and Noray (2018).

[5] While an exploration of the mechanisms that drive task changes is beyond the scope of this paper, we explore one such mechanism in related work (Atalay et al. 2018). In that paper, we extract additional information from vacancy postings: mentions of 48 distinct information and communication technologies. Based on the patterns of task and technology mentions, we argue that technologies tend to increase the demand for worker-performed nonroutine analytic tasks relative to other tasks (though exceptions, like the Microsoft Office Suite, exist).

[6] Gentzkow, Kelly, and Taddy (2019) summarize recent applications of text analysis in economic research.

[7] Our paper also relates to work by Abraham and Wachter (1987). Using the Help Wanted Index, they document that frequencies of newspaper job ad postings track reasonably well with administrative data on total labor market vacancies. A more recent paper that uses newspaper job ads is DeVaro and Gürtler (2018), which studies

literature is to extend the analysis of job ad text to the pre-internet era, spanning a much longer horizon and a key period of occupational change. We also apply tools from natural language processing—which to our knowledge have had limited use in economics research—to extend our word-based task categories to include synonyms for task-related words and to limit the sensitivity of our analysis to changes in word meaning over time.

The rest of the paper is organized as follows. Section I outlines the construction of our dataset of occupations and their content, then compares this new dataset to existing data sources. In Section II, we document changes in occupational tasks in the aggregate and conclude that a large share has occurred within, rather than between, occupations. Section III provides new descriptive evidence on the changing nature of work through a discussion of a few selected job titles. Section IV concludes and suggests areas for future research.

I. A New Dataset of Occupational Characteristics

In this section, we discuss the construction of our structured database of occupational characteristics. The primary datasets are raw text files purchased from ProQuest and were originally published in the *New York Times* (from 1940 to 2000), *Wall Street Journal* (from 1940 to 1998), and *Boston Globe* (from 1960 to 1983).[8] The first step in our approach is to clean and process the raw newspaper data. We then map job ad titles to SOC codes and map job ad text into task categories. Once we describe these procedures, we illustrate the performance of our approach using a set of ads from the April 10, 1960 *New York Times* and present some simple descriptive statistics from our dataset. Lastly, we describe several checks we perform on the data to test for selection and time-varying measurement error.

A. *Processing the Newspaper Text Files*

The newspaper data are stored as raw text files, which ProQuest has produced using an algorithm that converts images of newspapers into text files. The raw text files ProQuest has provided allow us to isolate the subset of text that comes from advertisements, but do not allow us to directly identify job ads from other types of advertisements. The text also does not indicate where one job ad ends and another begins. Therefore, in processing the ProQuest text files, we must (i) identify which advertisements comprise vacancy postings, (ii) discern the boundaries between vacancy postings, and (iii) identify the job title of each vacancy posting. In addition, as much as possible, we attempt to undo the spelling mistakes induced by ProQuest's imperfect transcription of the newspaper text. Online Appendix D.1 describes our procedure for performing (i). Online Appendix D.2

worker-firm matching. They document that before 1940, both job seekers and firms posted advertisements to match with one another. Since 1940, firms have been the primary party posting ads.

[8] In addition to the newspaper text files, we use a dataset purchased from Economic Modeling Specialists International (EMSI); these data include the full text of the near-universe of online job ads for selected months between 2012 and 2017. As described below, we use the EMSI data to identify word synonyms and to study the geographic selection of job ads.

describes steps (ii) and (iii). Overall, our procedure allows us to transform unstructured text into a set of 8.3 million distinct job ads linking job titles to job ad text.[9]

B. *Grouping Occupations by SOC Code*

Our next step is to consolidate the information in our vacancy postings to characterize occupations and their corresponding attributes into a small number of economically meaningful categories. In the newspaper text, postings for the same occupation appear via multiple distinct job titles. For example, vacancy postings for registered nurses will be advertised using job titles that include "IV nurse," "ICU nurse," or "RN coordinator." These job titles should all map to the same occupation code 291141, using the SOC system.

From our list of job titles, we apply a *continuous bag of words* (CBOW) model to identify the ad's SOC code. Roughly put, this CBOW model allows us to find synonyms for words or phrases. The model is based on the idea that words or phrases are similar if they themselves appear (in text corpora) near similar words. For example, to the extent that "IV nurse," "ICU nurse," and "RN coordinator" all tend to appear next to words like "patient," "care," or "blood," one would conclude that "RN" and "nurse" have similar meanings. For additional background on CBOW models and details of our implementation, see online Appendix D.3. EMSI has provided us with a dataset of the text from online job ads originally posted between October 2011 and March 2017. These ads contain a job title and text describing the job characteristics and requirements. We use online job postings from two of these months, January 2012 and January 2016, plus all of the text from our newspaper data to construct our CBOW model.

Our CBOW model is useful for our purposes when applied in combination with O*NET's Sample of Reported Titles and list of Alternate Titles. Once we have estimated the CBOW model, for each job title $\mathcal{N}$ in our newspaper text, we search for the job title $\mathcal{O}$ among those in the O*NET Sample of Reported Titles and list of Alternate Titles that is most similar to $\mathcal{N}$.[10] Since each of the job titles in the O*NET Sample of Reported Titles and list of Alternate Titles has an associated SOC code, we can obtain the SOC code for any job title in our newspaper text. As an example, the job title "RN coordinator"—a title from our newspaper data—is closest to the O*NET Title "Registered Nurse Supervisor," which has an associated SOC code of 291141. Based on this, we identify 291141 as the SOC code for "RN coordinator." In this manner, we retrieve these SOC

[9] This 8.3 million figure excludes vacancy postings for which we cannot identify the job title or that contain a substantial portion (35 percent or greater) of misspelled words. We also exclude ads with fewer than 15 words.

[10] The CBOW model associates each word and phrase with a vector, with elements in the vector describing the contexts in which the word or phrase appears. The similarity between job titles $\mathcal{O}$ and $\mathcal{N}$ equals the cosine similarity of the vectors associated with these two titles.

codes on all of the job titles that appear in our newspaper text.[11] This procedure yields SOC codes for 7.8 million job ads.[12]

C. *Eliciting Job-Related Information*

Within the body of our job ads, we map similar words to a common task or skill. For example, mathematical skills could appear in job ads using the words "mathematics," "math," or "quantitative." To study occupations' evolving skill requirements and task content, it is necessary to categorize these occupational characteristics into a manageable number of groups. We follow three approaches, which we explain next.[13]

Our main classification follows that of Spitz-Oener (2006), who in her study of the changing task content of German occupations, groups survey questionnaire responses into five categories: *nonroutine analytic*, *nonroutine interactive*, *nonroutine manual*, *routine cognitive*, and *routine manual*.[14] In our main application of these categories, we begin with the list of words related to each of her five tasks. For each task, we augment the list with words that have similar meanings to those in the original list, where similarity is determined by the same CBOW model introduced in Section IB. This is our primary classification, and we use it in each empirical exercise that follows in the paper. In addition, as a robustness check, we consider a narrower mapping between categories and words, one that only relies on Spitz-Oener's (2006) definitions as enumerated in footnote 14. Including similar words based on our CBOW model has its advantages and disadvantages. On the one hand, the CBOW model has the advantage of accounting for the possibility that the word choice of employers may differ within the sample period.[15] On the other

[11] For our 1950–2000 sample period, we cannot directly evaluate the accuracy of our SOC assignment algorithm. However, the online job ad data we have procured from EMSI contain an SOC code, which allows us to assess the performance of our method to assign SOC codes in a more recent dataset. To do so, we compare the results from our procedure to the SOC code available in the EMSI data. Our procedure assigns the same four-digit SOC code 53 percent of the time and the same six-digit SOC code 36 percent of the time. As there are 110 unique four-digit SOC codes and 840 unique six-digit SOC codes, these rates suggest our algorithm has a high degree of precision.

[12] We do not find an associated SOC code for certain job titles, such as "trainee" or "personnel secretary," for which the title is either uninformative (in the case of trainee) or refers to the person to whom job applications are usually sent (in the case of personnel secretary). For this reason, our main dataset includes fewer than the 8.3 million ads mentioned at the end of Section IA.

[13] Throughout this paper, we interpret the words as accurate representations of the positions firms seek to fill. We cannot measure the extent to which firms may misrepresent or perhaps euphemize the tasks of the job to attract workers. A similar consideration, however, is also relevant for survey-based measures of tasks, in which respondents may or may not accurately answer questions about their job's tasks (Autor 2013). Our analysis is unaffected by level differences in job descriptions' accuracy, and it would only be affected by trends in the representation of jobs over time.

[14] The dataset used by Spitz-Oener (2006) is a questionnaire given to West German workers. Building on her mapping from survey question titles to task categories, we search for the following sets of words for each category: (i) nonroutine analytic: analyze, analyzing, design, designing, devising rule, evaluate, evaluating, interpreting rule, plan, planning, research, researching, sketch, sketching; (ii) nonroutine interactive: advertise, advertising, advise, advising, buying, coordinate, coordinating, entertain, entertaining, lobby, lobbying, managing, negotiate, negotiating, organize, organizing, presentation, presentations, presenting, purchase, sell, selling, teaching; (iii) nonroutine manual: accommodate, accommodating, accommodation, renovate, renovating, repair, repairing, restore, restoring, serving; (iv) routine cognitive: bookkeeping, calculate, calculating, correcting, corrections, measurement, measuring; and (v) routine manual: control, controlling, equip, equipment, equipping, operate, operating.

[15] For instance, even though "creative" and "innovative" largely refer to the same occupational skill, it is possible that their relative usage among potential employers may differ within the sample period. This is indeed

hand, there is a danger that the CBOW model will identify words as synonymous even if they are not.

We also consider alternative and complementary task classifications for the purpose of (i) exploring the robustness of our results to our primary choice of classification, (ii) comparing our text-based measures with widely used survey-based measures, and (iii) connecting our main results to those in the literature. Our second classification draws on the groups of skills that Deming and Kahn (2018) have defined in their study of the relationship between the characteristics of firms and the skill requirements in their vacancy postings.[16] Finally, with the aim of validating our dataset, we map our text to O*NET's work styles, skills, knowledge requirements, and work activities (corresponding to O*NET Elements 1C, 2A and 2B, 2C, and 4A, respectively). As with our Spitz-Oener-based measures, we append synonymous words—using our CBOW model—to the lists of skill-related words and phrases of Deming and Kahn (2018) and O*NET.

D. *An Example from the April 10, 1960 New York Times*

Having delineated our procedure for cleaning and extracting information from our newspaper text, we next illustrate the performance of our procedure with an example. Figure 1 presents a snippet of digitized text from a page of display ads in the April 10, 1960 edition of the *New York Times*. This text refers to multiple vacancy postings, including one for an accountant position, a second for a mechanical engineer position, a third for a methods engineer position, and so on. Each ad describes, in varying levels of detail, the sets of tasks that workers will perform, experience requirements, and aspects of the work environment. Some, but not all, of the ads contain the identity of the posting firm. Some ads contain a posted salary, while others do not. As Figure 1 makes clear, while the text contains a high frequency of transcription errors due to the imperfect performance of ProQuest's optical character recognition technology, much of the information contained in the ad is preserved.

Figure 2 presents the output from our approach. First, on the basis of strings that tend to appear at the beginning and end of job ads, our algorithm successfully finds the boundary between the accountant and mechanical engineer job ad and between the transportation advertising supervisor and performance engineer

the case. Use of the word "innovative" has increased more quickly than "creative" over the sample period. If our classification included only one of these words, we would be mischaracterizing trends in the O*NET skill of "thinking creatively." The advantage of the continuous bag of words model is that it will identify "creative" and "innovative" as synonyms because they appear in similar contexts within job ads. Hence, even if employers start using "innovative" as opposed to "creative" partway through our sample, we will be able to consistently measure trends in "thinking creatively" throughout the entire period. A second advantage of our CBOW model is that it allows us to partially undo the transcription errors generated by ProQuest's image scanning. Our CBOW algorithm, for example, identifies "adverhslng" as synonymous with "advertising."

[16] See table 1 of Deming and Kahn (2018) for their list of words and their associated skills. Building on their definitions, we use the following rules: (i) cognitive: analytical, cognitive, critical thinking, math, problem solving, research, statistics; (ii) social: collaboration, communication, negotiation, presentation, social, teamwork; (iii) character: character, energetic, detail oriented, meeting deadlines, multitasking, time management; (iv) writing: writing; (v) customer service: client, customer, customer service, patient, sales; (vi) project management: project management; (vii) people management: leadership, mentoring, people management, staff, supervisory; (viii) financial: accounting, budgeting, cost, finance, financial; and (ix) computer (general): computer, software, spreadsheets.

TIMES ACCOUNTANTS Due to staff promotions, openings have developed in our Cost and Auditing Divisions of parent company. We are looking for men with 2 to 5 years of experience with a large public accounting firm. Good opportunities for growth. Excellent salary. Send resume to Personnel Department Johnson &Johnson. New Brunswick, New Jersey MECHANICAL ENGINEER Specialist In selection of pumps, compressors &general mechanical equipment. 4 te 6 yrs exp. with pump mfr.. engineering contractor. o or public utility, etc. o . . Good starting salary o . Ercellent conditions ark Area BOX 219, Large New England sheet metal fabricating plant manufacturing extensive line of InstItutional furniture has good opportunity for Methods Engineer with comprehensive knowledge of ope,ations and layout. Include resume and salary. requirements. X7548 TIMES u RESUMES PRINTED S3.50 lst5Goiviesfnciudiiw type. Si ch - add. 100 copIes. I Add 35c to mall ord (P1AE) Open DiSh td6 P.M. DAY The PRESS 42Wust 33 SI4E6Y.C. OX 5.3658 Major Oil Company Needs A TRANSPORTATION ADVERTISING SUPERVISOR With Specific experience in creating advertising for: truck-bus, aviation, marine or construction industries. Understanding of advertising media, creative functions, agency relationships and organization procedures. College degree with a background in advertising and sales promotion. Versatility, initiative and a good personality. Some knowledge of the, petroleum requirements and their application to the transportation industries desirable. OPPORTUNITY FOR ADVANCEMENT ? by letter only, submitting detailed resume of education, experience and salary requirements. Socony Mobil Oil Company, Inc. 150 East 42 Street, N. Y. (at Lexington) PERFORMANCE ENGINEERS Aircraft &Space Vehicle Systems Evaluation Diversified projects include the evaluation of advanced propulsion concepts for subsonic, hypersonic and space vehicles in terms of system performance capabilities. Sustained program with excellent support from services from the largest industrial computing efforts by experienced component specialists. Minimum qualifications for these positions include a M.S. degree in aeronautical engineering plus 3 related experience. UNITED AIRCRAFT CORPORATION 400 Main Street . East Hartford, Conn. Please write to Mr. W. M. Walsh RESEARCH LABORATORIES

FIGURE 1. UNPROCESSED ADS FROM THE APRIL 10, 1960 *NEW YORK TIMES*

Notes: This figure presents the digitized text, obtained from ProQuest, from a portion of page F13 of the April 10, 1960 edition of the *New York Times*. In online Appendix D.2, we present the corresponding partial page of ads as it originally appeared.

job ad. However, some of the text from the Methods Engineer ad is erroneously appended to that of the preceding ad.

Further, our algorithm extracts task-related words. The transportation advertising supervisor ad includes three mentions of nonroutine interactive tasks—"advertising," "media," and "sales"—and a single mention of a nonroutine analytic task—"creating." Of the task words in Spitz-Oener (2006), only "advertising" appears in the raw text presented in Figure 1. The others are included on the basis of our CBOW algorithm. This algorithm identifies "media" as close to "advertising," "sales" as close to "selling," "creating" as close to "designing," and "evaluation" as

ACCOUNTANT [[132011]] Due to staff promotions, openings have developed in our Cost and Auditing Divisions of parent company. We are looking for men with 2 to 5 years of experience with a large public accounting firm. Good opportunities for growth. Excellent salary. Send resume to Personnel Department Johnson &Johnson. New Brunswick, New Jersey

MECHANICAL ENGINEER [[172141]] Specialist In selection of pumps, compressors &general mechanical equipment. 4 te 6 yrs exp. with pump mfr. engineering contractor. o or public utility, etc. o . . Good starting salary o . Ercellent conditions ark Area BOX 219, Large New England sheet metal fabricating plant manufacturing extensive line of InstItutional furniture has good opportunity for Methods Engineer with comprehensive knowledge of ope,ations and layout. Include resume and salary. requirements. X7548 TIMES u RESUMES PRINTED S3.50 lst5Goiviesfnciudiiw type. Si ch - add. 100 copIes. I Add 35c to mall ord (P1AE) Open DiSh td6 P.M. DAY The PRESS 42Wust 33 SI4E6Y.C. OX 5.3658 Major Oil Company Needs A

TRANSPORTATION ADVERTISING SUPERVISOR [[531031]] With Specific experience in creating advertising for: truck-bus, aviation, marine or construction industries. Understanding of advertising media, creative functions, agency relationships and organization procedures. College degree with a background in advertising and sales promotion. Versatility, initiative and a good personality. Some knowledge of the, petroleum requirements and their application to the transportation industries desirable. OPPORTUNITY FOR ADVANCEMENT ? by letter only, submitting detailed resume of education, experience and salary requirements. Socony Mobil Oil Company, Inc. 150 East 42 Street, N. Y. (at Lexington)

PERFORMANCE ENGINEER [[173029]] Aircraft &Space Vehicle Systems Evaluation Diversified projects, include the evaluation of advanced propulsion concepts for subsonic, hypersonic and space vehicles in terms of system performance capabilities. Sustained program with excellent support from services from the largest industrial computing efforts by experienced component specialists. Minimum qualifications for these positions include a M.S. degree in aeronautical engineering, plus 3 related experience.
UNITED AIRCRAFT CORPORATION 400 Main Street . East Hartford, Conn. Please write to Mr. W. M. Walsh RESEARCH LABORATORIES

Figure 2. Processed Ads from the April 10, 1960 *New York Times*

Notes: We identify four ads from the unprocessed text in Figure 1. The job titles that we have identified, located at the beginning of each ad, are written in bold. Using Spitz-Oener's (2006) lists of task categories, we highlight nonroutine analytic tasks: "creating" and "evaluation" refer to nonroutine analytic tasks. We place in rectangles words that refer to nonroutine interactive tasks: "advertising," "media," and "sales." We also search for nonroutine manual, routine cognitive, and routine manual tasks. There are no mentions of these tasks within these ads. Using Deming and Kahn's (2018) lists, we place in ovals words that refer to financial skills: "accounting," "auditing," and "cost." We place in dashed trapezoids words that refer to project management: "construction," "projects," and "engineering." In addition, "computing" in the performance engineer ad refers to Deming and Kahn's (2018) computer skill; "sales" in the transportation advertising supervisor ad refers to the customer service skill; and "staff" and "personnel" in the accountant ad refer to the people-management skill. There were no mentions of other Deming and Kahn (2018) skills in these four ads. The six-digit code in square brackets refers to the SOC code we have identified: 132011 is the code for accountants and auditors, 172141 is the code for mechanical engineers, 531031 is the code for first-line supervisors of transportation and material-moving machine and vehicle operators, and 173029 is the code for engineering technicians.

close to "evaluating." Within the figure, we also highlight words corresponding to Deming and Kahn's (2018) lists of skills: "cost," "auditing," and "accounting" correspond to their financial skill category; "engineering," "construction," and "projects" correspond to their project management skill; and so on.

Finally, our algorithm reasonably identifies SOC codes for each job title. It assigns the accountant job title to the SOC code for accountants and auditors, the mechanical engineer job title to the SOC code for mechanical engineers, and the performance engineer job title to the SOC code for other engineering technicians. The transportation advertising supervisor job title, which we have classified in the first-line supervisors of transportation operators SOC, presents an ambiguous case; it could reasonably be classified among advertising-related or transportation-related occupations. Overall, our text-processing procedure satisfactorily retrieves information on the titles of jobs, their associated tasks, and their occupational codes.

TABLE 1—COMMON OCCUPATIONS

Job title		Six-digit SOC occupations		Four-digit SOC occupations	
Description	Count	Description	Count	Description	Count
Secretary	165.0	439022: Typist	363.3	4360: Secretary	638.8
Typist	103.5	436012: Legal secretary	266.4	4390: Other admin.	564.1
Clerk	97.9	414012: Sales rep.	239.2	4330: Financial clerks	363.2
Assistant	89.4	412031: Retail sales	234.6	1320: Accountant	344.5
Sales	86.5	132011: Accountant	217.3	4140: Sales rep.	286.0
Salesperson	83.9	436014: Secretary	199.8	4120: Retail sales	278.1
Bookkeeper	69.6	436011: Exec. secretary	149.7	4340: Record clerks	227.3
Accounting	69.5	433031: Bill collectors	148.6	1511: Computer Sci.	202.6
Clerk typist	68.5	434031: Credit authorizers	145.1	1720: Engineers	199.5
Engineer	62.8	433021: Bookkeeper	138.0	1730: Drafters	178.6

Notes: This table lists the top 10 job titles (columns 1–2), the top 10 six-digit SOC codes (columns 3–4), and the top 10 four-digit SOC codes (columns 5–6) in the *Boston Globe*, *New York Times*, and *Wall Street Journal*. Counts are given in thousands of newspaper job ads.

E. *Descriptive Statistics*

Using the newspaper text, our algorithm from Sections IA, IB, and IC results in a dataset with 7.8 million vacancy postings. Table 1 lists the top occupations in our dataset. The first two columns list the most common job titles, and the last four columns present the most common SOC codes.[17] Across the universe of occupations, our newspaper data represent a broad swath of management, business, computer, engineering, life and physical science, healthcare, sales, and administrative support occupations, but they underrepresent construction occupations and occupations related to the production and transportation of goods.

To systematically assess the representativeness of our newspaper data, in online Appendix B.1, we compare the share of workers across occupations in the decennial census to the share of vacancies in our new dataset (Ruggles et al. 2015). Perhaps unsurprisingly, our newspaper text underrepresents certain blue-collar occupations. Nevertheless, there are still a considerable number of ads that we can map to each six-digit SOC code throughout our sample period, including broad coverage of blue-collar occupations. In the same online Appendix, we establish that occupations the decennial census measures as having a large share of educated workers also tend to have newspaper ads with a large share of stated education requirements.[18]

Table 2 presents, for each task in Spitz-Oener's (2006) classification, the most task-intensive occupations. For each job title-year combination, we first compute the number of mentions of task h per 1,000 job ad words, $\tilde{T}_{j,t}^{h}$, and the fraction of year t ads that have j as the job title, $S_{j,t}$. Then, for each of the 200 most commonly appearing job titles, we compute the average of $\tilde{T}_{j,t}^{h}$ (and also $S_{j,t}$) across the years in our sample. We find that engineering jobs are among the occupations most intensive

[17] Marinescu and Wolthoff (forthcoming) document that many job titles contain multiple words. Even though the top job titles in Table 1 are single word, most job ads—73 percent—contain multiword job titles. To the extent that newspaper space is scarcer than space within online job ads, newspaper job titles will be shorter than the job ads within Marinescu and Wolthoff's (forthcoming) analysis.

[18] We also consider the distribution of vacancies across occupations in our data compared to employment shares in Boston and New York in the decennial census. Not surprisingly, our vacancy data more closely track the occupational shares in Boston and New York than those in the United States as a whole, but they track US employment shares notably well.

Table 2—Top Job Titles by Spitz-Oener (2006) Task Category

Nonroutine analytic			Nonroutine interactive		
Design engineer	0.0005	20.06	Sales manager	0.0018	18.50
Mechanical engineer	0.0010	19.28	Account executive	0.0010	17.53
Systems engineer	0.0004	18.81	Sales executive	0.0006	16.92
Electrical engineer	0.0007	17.92	Sales representative	0.0017	15.88
Project engineer	0.0006	16.87	Sales engineer	0.0012	15.61
Nonroutine manual			Routine cognitive		
Mechanic	0.0014	6.58	Payroll clerk	0.0008	15.90
Auto mechanic	0.0006	4.97	Billing clerk	0.0005	13.74
Electronic technician	0.0007	4.37	Bookkeeper full charge	0.0015	12.09
Electrician	0.0006	4.33	Assistant bookkeeper	0.0022	11.37
Superintendent	0.0016	3.75	Bookkeeper	0.0083	9.76
Routine manual					
Machinist	0.0012	5.29			
Mechanic	0.0014	2.45			
Mechanical engineer	0.0010	1.77			
Foreman	0.0016	1.64			
Design engineer	0.0005	1.35			

Notes: This table lists the top five job titles according to the frequency with which different activity-related words are mentioned. Within each panel, the first column gives the job title; the second column gives $1/51 \cdot \sum_{t=1950}^{2000} S_{j,t}$—i.e., the average share of ads belonging to the job title; and the final column gives the average frequency of task h words among job title j's ads per 1,000 ad words.

in nonroutine analytic tasks. Sales occupations mention nonroutine interactive tasks most frequently, while mechanical and electrical occupations rank highest in their intensity of nonroutine manual tasks. Clerical and production-related positions mention routine cognitive and routine manual task-related words most frequently.

F. *Comparison to Existing Datasets*, *Selection*, *and Time-varying Measurement Error*

Thus far we have shown that our data contain valuable information about job tasks. Before our main analysis, we compare our new dataset to existing data sources. We then discuss the robustness checks we perform on the data to test for changing patterns of selection or measurement error over the sample period.

First, we compare occupation-specific measurements in our dataset to those in existing data sources. In online Appendix B.2, we compare occupations' O*NET importance scores—for various O*NET work styles, skill requirements, knowledge requirements, and work activities—to the frequency of words corresponding to these job attributes in our new dataset. We find that cross-sectional correlations, looking across occupations, between O*NET's measures and measures from our newspaper data fall for the most part within the 0.40 to 0.65 range. In online Appendix B.3, we compare our new dataset with the DOT. We show that across occupations, tasks are positively correlated with those measured in the DOT. We then consider the DOT's usefulness for time series analysis and argue that while carefully constructed in the cross section, these measures miss much of the evolution of occupational characteristics over time.

Second, in online Appendix C.1, we document that there is no more than a weak, marginally significant trend in average ad length and that there are no meaningful trends in the share of ad words that do not appear in an English dictionary. Such trends, if they were present and important in the data, would be suggestive of trends in measurement error within our sample period (due, for example, to changing typographical conventions or improvements in image quality). Based on these exercises, we do not find evidence of time-varying measurement error. We also emphasize that because we use the CBOW model, our analysis is less sensitive to changing diction over time, since word substitutions of synonyms would be measured uniformly as task words.

In online Appendix C.2, we consider the possibility of trends in selection of posting vacancies in newspapers over time. While we do not observe the firm's decision directly, we are able to use the Current Population Survey to study the worker side and the decision to search using alternative methods. Specifically, we check whether workers exhibit trends in their propensity to search using jobs ads; we also investigate whether there are differential trends by the task intensity of their prior occupation. If, for example, workers in occupations that are high in nonroutine tasks are more likely to search in newspapers over time compared to workers in occupations low in nonroutine tasks, we would be concerned that selection is causing us to overstate the upward trend in nonroutine tasks. The analysis in online Appendix C.2 shows that there are no differential trends in selection over time on the worker search side, which provides some suggestive evidence that selection into newspaper posting has been stable over our period of analysis.

There is still the possibility of geographic selection *within* occupations over time. Without an available dataset to measure within-occupation tasks over the sample period, we cannot directly test for this source of bias. However, we can perform a direct test over a more recent, out-of-sample period, which we do in online Appendix C.3. Using a 5 percent sample of ads that were posted online and collected by EMSI (totaling 7.6 million ads), we compare the task content of ads that were posted for jobs in the New York City and Boston metro areas to jobs based elsewhere. In particular, we examine whether within occupations, the task content is systematically different in Boston and New York City compared to the rest of the United States, and whether there have been differential trends in occupational tasks. We find that job ads in the New York City and Boston metro areas are indeed statistically different but only slightly so; the difference amounts to at most 0.05 standard deviations of each of these task measures. We find that these differences are even smaller within occupations. Overall, geographic selection by task intensity appears to be minor when compared to the overall dispersion in task measures across all online job ads.

II. Trends in Tasks

In this section, we document trends in occupational tasks from 1950 to 2000. We show first that the labor market has experienced dramatic shifts toward nonroutine tasks and away from routine ones. Second, we show that a large part of this change reflects an evolution of occupations themselves, with a smaller fraction accounted for by shifts in employment across occupations. We emphasize that this finding holds regardless of how finely we define occupations or whether we use the Spitz-Oener or the Deming and Kahn classification. Finally, we argue that these findings, while demonstrating that the transformation of the US workplace has been larger than previously thought, are entirely consistent with previous findings in the literature and in particular with those of Autor, Levy, and Murnane (2003).

A. *Overall Trends*

Table 3 presents changes in task mentions, grouped according to the definitions introduced in Spitz-Oener (2006). In each of the five panels within this table, the first row presents the mean task frequency per 1,000 words at the beginning of the sample, in 1950. We call this $\overline{T}_{1950}$. In 1950, the economy-wide average was 2.77 mentions of nonroutine analytic tasks, 5.06 mentions of nonroutine interactive tasks, 0.91 mentions of nonroutine manual tasks, 1.89 mentions of routine cognitive tasks, and 0.97 mentions of routine manual tasks. In the remaining rows of each panel, we display changes in the task mentions. The measure in the first column of each panel, which we interpret as the aggregate task content in the economy, presents changes in task mentions across occupations. According to this measure, the frequency of nonroutine analytic tasks increased by 75 log points between 1950 and 2000. Over the same period, the frequency of nonroutine interactive tasks increased by 38 log points. Conversely, routine manual task mentions substantially declined, decreasing from 0.97 to 0.06 mentions per 1,000 job ad words. The decline of routine cognitive tasks is also considerable, going from 1.89 to 0.85 mentions per 1,000 job ad words.

These changes reflect both between-occupation and within-occupation changes in tasks. To assess the relative importance of between- versus within-occupation forces in shaping these trends, we decompose changes in the aggregate content of each task according to the following equation:

$$\overline{T}_t = \overline{T}_{1950} + \sum_j \vartheta_{j,1950}\left(\tilde{T}_{j,t} - \tilde{T}_{j,1950}\right) + \sum_j \left(\vartheta_{j,t} - \vartheta_{j,1950}\right)\tilde{T}_{j,t}. \tag{1}$$

In this equation, $\tilde{T}_{j,t}$ measures the frequency of task-related words for occupation j in year t.[19] The $\vartheta_{j,t}$ terms measure the share of workers in occupation j at time t according to the decennial census, while $\overline{T}_t$ denotes the average frequency of the task-related word at time t.[20] On the right-hand side of equation (1), the first

[19] In Table 3 and in subsequent decomposition tables, each decade t contains surrounding years to reduce the effect of sampling error: "1950" contains ads from 1950 to 1953, "1960" contains ads from 1958 to 1962, "1970" contains ads from 1968 to 1972, "1980" contains ads from 1978 to 1982, "1990" contains ads from 1988 to 1992, and "2000" contains ads from 1997 to 2000.

[20] Throughout this section, we draw from the sample of full-time workers—workers who are between the ages of 16 and 65; who work for wages; who worked at least 40 weeks in the preceding year; and who have nonimputed gender, age, occupation, and education data. We construct our own mapping between four-digit SOC codes and census occ1990 codes by taking the modal SOC code for each occ1990 code (drawing on a sample of all workers in the 2000 census public use sample and the 2007 and 2013 American Community Survey for which both variables are measured). From our full-time worker sample, we compute the share of workers who work in four-digit SOC occupation o in decennial census years; call this number $\vartheta_{o^4,t}$. Then, to compute weights for six-digit SOC code occupations, we multiply $\vartheta_{o^4,t}$ by the fraction of year-t ads for the six-digit SOC code (within j's four-digit SOC). Below, when we use j to refer to a job title, to compute $\vartheta_{j,t}$, we multiply $\vartheta_{o^4,t}$ by the fraction of year-t ads (within j's four-digit SOC code) that correspond to job title j.

Table 3—Trends in Keyword Frequencies: Six-digit SOCs

	Total	Within	Between	Within share	Total	Within	Between	Within share
	Panel A. Nonroutine analytic				*Panel B. Nonroutine interactive*			
1950 level	2.77				5.06			
	(0.03)				(0.07)			
1950–1960	0.59	0.31	0.28	0.53	−0.11	−0.17	0.06	1.50
	(0.06)	(0.06)	(0.01)	(0.05)	(0.07)	(0.07)	(0.02)	(0.92)
1960–1970	−0.07	−0.17	0.10	2.40	−0.44	−0.60	0.16	1.37
	(0.07)	(0.06)	(0.02)	(15.32)	(0.05)	(0.06)	(0.02)	(0.07)
1970–1980	0.96	0.67	0.29	0.70	1.12	0.84	0.28	0.75
	(0.08)	(0.07)	(0.04)	(0.03)	(0.06)	(0.07)	(0.04)	(0.04)
1980–1990	0.28	0.20	0.08	0.72	0.98	0.99	−0.00	1.01
	(0.10)	(0.10)	(0.07)	(0.27)	(0.08)	(0.12)	(0.08)	(0.09)
1990–2000	1.35	1.69	−0.34	1.25	0.79	0.66	0.13	0.84
	(0.10)	(0.18)	(0.14)	(0.11)	(0.10)	(0.17)	(0.14)	(0.19)
1950–2000	3.11	2.70	0.40	0.87	2.33	1.72	0.62	0.74
	(0.07)	(0.16)	(0.13)	(0.04)	(0.11)	(0.17)	(0.09)	(0.05)
	Panel C. Nonroutine manual				*Panel D. Routine cognitive*			
1950 level	0.91				1.89			
	(0.03)				(0.03)			
1950–1960	−0.09	−0.09	0.00	1.03	−0.69	−0.63	−0.07	0.90
	(0.05)	(0.05)	(0.01)	(0.40)	(0.03)	(0.03)	(0.01)	(0.01)
1960–1970	−0.06	−0.09	0.03	1.57	−0.25	−0.32	0.07	1.29
	(0.03)	(0.03)	(0.01)	(1.70)	(0.03)	(0.03)	(0.01)	(0.05)
1970–1980	0.32	0.41	−0.09	1.29	−0.11	−0.11	−0.00	0.97
	(0.02)	(0.03)	(0.01)	(0.04)	(0.02)	(0.02)	(0.01)	(0.11)
1980–1990	−0.33	−0.37	0.04	1.14	0.03	0.01	0.02	0.30
	(0.03)	(0.03)	(0.02)	(0.06)	(0.03)	(0.03)	(0.02)	(1.99)
1990–2000	−0.01	−0.02	0.01	3.29	−0.02	−0.02	−0.00	0.98
	(0.04)	(0.04)	(0.03)	(2.95)	(0.04)	(0.04)	(0.03)	(3.20)
1950–2000	−0.16	−0.17	0.00	1.03	−1.04	−1.06	0.02	1.02
	(0.04)	(0.04)	(0.02)	(0.15)	(0.03)	(0.04)	(0.03)	(0.03)
	Panel E. Routine manual							
1950 level	0.97							
	(0.03)							
1950–1960	−0.25	−0.21	−0.05	0.81				
	(0.03)	(0.04)	(0.01)	(0.05)				
1960–1970	−0.28	−0.29	0.01	1.03				
	(0.02)	(0.03)	(0.01)	(0.04)				
1970–1980	−0.03	0.12	−0.15	−3.70				
	(0.01)	(0.02)	(0.01)	(3.71)				
1980–1990	−0.27	−0.39	0.12	1.42				
	(0.01)	(0.03)	(0.02)	(0.05)				
1990–2000	−0.08	−0.01	−0.06	0.18				
	(0.01)	(0.05)	(0.05)	(0.67)				
1950–2000	−0.91	−0.78	−0.14	0.85				
	(0.03)	(0.06)	(0.05)	(0.05)				

Notes: Occupations are defined using the six-digit SOC classification. Within each panel, we compute keyword frequencies per 1,000 ad words at the beginning of the sample (first row), decade-by-decade changes (second through sixth rows), and cumulative changes over the 50-year period (seventh row). In these averages, occupation shares for each four-digit SOC are given by the number of full-time workers in the decennial census. Terms within parentheses give bootstrapped standard errors based on resampling ads from our newspaper text 40 times.

sum captures shifts in the overall mentions due to within-occupation changes in task-related word mentions. The second sum captures shifts in the share of workers across occupations. We use a six-digit SOC classification to perform this decomposition separately for each of the five tasks introduced by Spitz-Oener (2006). The second and third columns of Table 3 list changes in $\overline{T}_t$ due to the within and between components of equation (1). The final column gives the proportion of the overall changes in the task due to the "Within" components.

This table shows that a substantial portion of the changes in task content have occurred within rather than between six-digit SOC occupations: 87 percent of the increase in nonroutine analytic tasks and 74 percent of the increase in nonroutine interactive tasks are due to within-occupation rather than between-occupation shifts in tasks. Similarly, all of the decline in routine cognitive tasks and 85 percent of the decline in routine manual tasks are due to within-occupation task shifts. Moreover, the within-occupation shifts are the primary source of task changes not only when looking over the 50-year period, but also when looking within each decade. Note that the "Within share" need not be bounded between 0 and 1 if within-occupation and between-occupation shifts in task content move in opposite directions.[21] Summing across the five task groups, 88 percent of the overall task changes occurred within six-digit SOC codes.[22]

The extent to which between-occupation changes are responsible for overall changes in tasks is potentially sensitive to the coarseness of occupation definitions. If occupations are coarsely defined, one would tend to estimate that between-occupation changes are relatively unimportant. To gauge the sensitivity of our results to our definition of occupation, Table 4 performs the same decomposition, now using job titles instead of six-digit SOC codes as the occupational unit. This is the finest classification one could possibly apply when decomposing trends in keyword frequencies into between-occupation and within-occupation components. As in Table 3, there has been a substantial shift away from routine manual and nonroutine analytic tasks. Also similar to the previous decomposition, within-occupation shifts account for a large majority—83 percent for nonroutine analytic tasks and 91 percent for routine manual tasks—of the overall changes. Overall, summing across the five task groups, 88 percent of the economy-wide task changes have occurred within job titles.

In Table 5, we use Deming and Kahn's categorization of skills. Among these skills, computer, customer-service, and social skills have increased most starkly. For seven of the nine skills, with financial and problem-solving skills as the two exceptions, within-job title changes are the primary source of growth in mentions of skill-related words.

[21] The "Within share" reported in the final columns of Table 3 is largely consistent with table 5 of Spitz-Oener (2006). There, Spitz-Oener (2006) calculates that nearly all of the changes in West German task content between 1979 and 1999 occurred within rather than between occupations.

[22] We first compute the overall changes between 1950 and 2000, summing across the five task categories: $4.91 = |\log((3.11 + 2.77)/2.77)| + |\log((2.33 + 5.06)/5.06)| + |\log((0.91 - 0.17)/0.91)| + |\log((1.89 - 1.04)/1.89)| + |\log((0.97 - 0.91)/0.97)|$. Second, we compute the portion of those changes that arises from the within-occupation component: $4.31 = 0.87 \cdot |\log((3.11 + 2.77)/2.77)| + 0.74 \cdot |\log((2.33 + 5.06)/5.06)| + 1.03 \cdot |\log((0.91 - 0.17)/0.91)| + 1.02 \cdot |\log((1.89 - 1.04)/1.89)| + 0.85 \cdot |\log((0.97 - 0.91)/0.97)|$. Taking the ratio of the two sums yields our 88 percent figure. Below, we use this as our summary statistic as the contribution of the within-occupation margin to overall task changes.

TABLE 4—TRENDS IN KEYWORD FREQUENCIES: JOB TITLES

	Total	Within	Between	Within share	Total	Within	Between	Within share
	Panel A. Nonroutine analytic				*Panel B. Nonroutine interactive*			
1950 level	2.86				5.03			
	(0.03)				(0.07)			
1950–1960	0.54	−0.04	0.58	−0.08	−0.04	−0.23	0.20	6.60
	(0.07)	(0.05)	(0.04)	(0.11)	(0.08)	(0.08)	(0.05)	(27.23)
1960–1970	−0.15	−0.06	−0.09	0.41	−0.53	−0.57	0.04	1.07
	(0.07)	(0.07)	(0.06)	(0.56)	(0.06)	(0.11)	(0.09)	(0.18)
1970–1980	0.67	0.30	0.37	0.45	1.08	0.63	0.44	0.59
	(0.07)	(0.07)	(0.08)	(0.10)	(0.06)	(0.12)	(0.10)	(0.10)
1980–1990	0.30	0.38	−0.08	1.28	0.99	1.03	−0.04	1.04
	(0.08)	(0.13)	(0.12)	(0.51)	(0.09)	(0.20)	(0.21)	(0.21)
1990–2000	1.26	1.58	−0.32	1.25	0.61	0.91	−0.30	1.49
	(0.10)	(0.29)	(0.28)	(0.22)	(0.10)	(0.31)	(0.34)	(0.73)
1950–2000	2.62	2.16	0.46	0.83	2.11	1.77	0.34	0.84
	(0.06)	(0.28)	(0.26)	(0.10)	(0.12)	(0.29)	(0.30)	(0.14)
	Panel C. Nonroutine manual				*Panel D. Routine cognitive*			
1950 level	0.97				1.99			
	(0.03)				(0.03)			
1950–1960	−0.10	−0.15	0.06	1.56	−0.72	−0.49	−0.22	0.69
	(0.05)	(0.05)	(0.02)	(4.23)	(0.03)	(0.04)	(0.02)	(0.04)
1960–1970	−0.07	−0.07	−0.00	0.96	−0.26	−0.32	0.06	1.21
	(0.04)	(0.03)	(0.02)	(0.38)	(0.03)	(0.04)	(0.04)	(0.15)
1970–1980	0.33	0.30	0.03	0.92	−0.13	−0.12	−0.01	0.93
	(0.02)	(0.04)	(0.03)	(0.10)	(0.02)	(0.04)	(0.04)	(0.35)
1980–1990	−0.34	−0.26	−0.09	0.75	0.05	0.14	−0.08	2.53
	(0.03)	(0.05)	(0.05)	(0.16)	(0.03)	(0.06)	(0.05)	(6.24)
1990–2000	−0.03	−0.01	−0.02	0.28	−0.04	−0.14	0.10	3.41
	(0.04)	(0.08)	(0.07)	(9.81)	(0.04)	(0.09)	(0.08)	(35.48)
1950–2000	−0.21	−0.19	−0.03	0.88	−1.10	−0.94	−0.16	0.86
	(0.04)	(0.06)	(0.05)	(0.26)	(0.03)	(0.07)	(0.07)	(0.06)
	Panel E. Routine manual							
1950 level	0.91							
	(0.04)							
1950–1960	−0.26	−0.20	−0.06	0.78				
	(0.04)	(0.04)	(0.02)	(0.08)				
1960–1970	−0.24	−0.25	0.02	1.07				
	(0.02)	(0.03)	(0.02)	(0.09)				
1970–1980	−0.02	0.11	−0.13	−5.88				
	(0.01)	(0.03)	(0.03)	(31.45)				
1980–1990	−0.27	−0.36	0.09	1.35				
	(0.01)	(0.03)	(0.03)	(0.10)				
1990–2000	−0.07	−0.07	0.00	1.02				
	(0.01)	(0.03)	(0.03)	(0.40)				
1950–2000	−0.85	−0.77	−0.08	0.91				
	(0.04)	(0.04)	(0.02)	(0.03)				

Notes: See the notes for Table 3. In comparison, we apply an occupation classification scheme based on job titles as opposed to six-digit SOC codes.

Table 5—Trends in Keyword Frequencies: Deming and Kahn (2018) Task Measures

	Total	Within	Between	Within share	Total	Within	Between	Within share
	Panel A. Character				*Panel B. Computer*			
1950 level	4.48				0.41			
	(0.07)				(0.01)			
1950–1960	−0.08	−0.01	−0.07	0.11	0.73	0.38	0.35	0.52
	(0.08)	(0.10)	(0.05)	(16.01)	(0.03)	(0.03)	(0.03)	(0.03)
1960–1970	0.80	0.53	0.27	0.66	0.19	0.24	−0.04	1.22
	(0.06)	(0.09)	(0.08)	(0.10)	(0.03)	(0.04)	(0.04)	(0.25)
1970–1980	1.60	1.77	−0.18	1.11	0.89	0.71	0.17	0.80
	(0.10)	(0.13)	(0.11)	(0.07)	(0.04)	(0.06)	(0.07)	(0.07)
1980–1990	0.60	0.51	0.09	0.84	1.29	1.75	−0.47	1.36
	(0.15)	(0.24)	(0.20)	(0.36)	(0.07)	(0.14)	(0.15)	(0.12)
1990–2000	−1.18	−1.26	0.08	1.07	1.37	1.11	0.26	0.81
	(0.10)	(0.26)	(0.24)	(0.20)	(0.09)	(0.17)	(0.17)	(0.12)
1950–2000	1.74	1.55	0.19	0.89	4.46	4.19	0.27	0.94
	(0.09)	(0.22)	(0.22)	(0.13)	(0.07)	(0.13)	(0.12)	(0.03)
	Panel C. Customer service				*Panel D. Financial*			
1950 level	2.86				2.45			
	(0.05)				(0.03)			
1950–1960	0.18	0.04	0.15	0.19	−0.26	−0.31	0.05	1.18
	(0.05)	(0.06)	(0.03)	(0.40)	(0.05)	(0.06)	(0.03)	(0.14)
1960–1970	−0.11	−0.11	−0.01	0.95	0.01	−0.27	0.28	−26.72
	(0.05)	(0.06)	(0.06)	(1.11)	(0.05)	(0.06)	(0.05)	(17.17)
1970–1980	0.82	0.82	−0.00	1.00	−0.03	−0.30	0.26	8.68
	(0.05)	(0.09)	(0.07)	(0.09)	(0.04)	(0.05)	(0.05)	(32.32)
1980–1990	1.65	1.45	0.20	0.88	0.35	0.31	0.04	0.89
	(0.07)	(0.15)	(0.13)	(0.08)	(0.04)	(0.04)	(0.05)	(0.14)
1990–2000	0.22	0.71	−0.49	3.21	0.29	0.29	0.00	0.98
	(0.08)	(0.26)	(0.26)	(2.83)	(0.05)	(0.11)	(0.12)	(0.44)
1950–2000	2.76	2.91	−0.15	1.05	0.35	−0.28	0.64	−0.80
	(0.09)	(0.27)	(0.25)	(0.09)	(0.05)	(0.12)	(0.10)	(0.42)
	Panel E. People management				*Panel F. Problem solving*			
1950 level	1.78				0.97			
	(0.02)				(0.02)			
1950–1960	0.59	0.30	0.29	0.51	0.35	0.11	0.24	0.33
	(0.04)	(0.04)	(0.04)	(0.07)	(0.04)	(0.04)	(0.03)	(0.10)
1960–1970	0.61	0.57	0.04	0.94	−0.31	−0.22	−0.09	0.71
	(0.04)	(0.06)	(0.06)	(0.10)	(0.03)	(0.03)	(0.03)	(0.10)
1970–1980	0.41	−0.03	0.44	−0.07	0.12	0.07	0.05	0.62
	(0.05)	(0.07)	(0.06)	(0.18)	(0.02)	(0.04)	(0.04)	(0.29)
1980–1990	0.09	0.20	−0.11	2.29	0.19	0.11	0.08	0.58
	(0.05)	(0.15)	(0.16)	(15.09)	(0.03)	(0.06)	(0.06)	(0.30)
1990–2000	−0.61	−0.36	−0.25	0.59	0.19	0.12	0.07	0.63
	(0.07)	(0.19)	(0.22)	(0.38)	(0.04)	(0.06)	(0.06)	(0.33)
1950–2000	1.09	0.68	0.41	0.63	0.55	0.20	0.34	0.37
	(0.05)	(0.11)	(0.12)	(0.10)	(0.03)	(0.06)	(0.04)	(0.09)

(*continued*)

Table 5—Trends in Keyword Frequencies: Deming and Kahn (2018) Task Measures (Continued)

	Total	Within	Between	Within share	Total	Within	Between	Within share
	Panel G. Project management				*Panel H. Social*			
1950 level	2.56				0.29			
	(0.03)				(0.01)			
1950–1960	0.82	0.01	0.80	0.02	0.07	−0.02	0.08	−0.24
	(0.07)	(0.06)	(0.05)	(0.07)	(0.02)	(0.01)	(0.01)	(0.30)
1960–1970	−0.21	−0.10	−0.11	0.47	0.03	0.04	−0.01	1.42
	(0.09)	(0.07)	(0.07)	(0.38)	(0.01)	(0.02)	(0.01)	(1.14)
1970–1980	0.87	0.38	0.50	0.43	0.39	0.27	0.12	0.69
	(0.10)	(0.08)	(0.09)	(0.08)	(0.01)	(0.03)	(0.03)	(0.07)
1980–1990	0.06	0.34	−0.28	6.16	0.70	0.64	0.06	0.91
	(0.10)	(0.12)	(0.11)	(19.64)	(0.03)	(0.05)	(0.05)	(0.07)
1990–2000	0.75	1.40	−0.65	1.87	0.41	0.46	−0.05	1.13
	(0.09)	(0.21)	(0.19)	(0.28)	(0.03)	(0.12)	(0.12)	(0.32)
1950–2000	2.28	2.03	0.25	0.89	1.59	1.38	0.21	0.87
	(0.07)	(0.17)	(0.15)	(0.07)	(0.03)	(0.12)	(0.12)	(0.07)
	Panel I. Writing							
1950 level	0.43							
	(0.01)							
1950–1960	0.02	−0.01	0.02	−0.30				
	(0.01)	(0.01)	(0.01)	(25.19)				
1960–1970	−0.13	−0.15	0.03	1.22				
	(0.01)	(0.02)	(0.01)	(0.11)				
1970–1980	0.11	0.12	−0.02	1.15				
	(0.01)	(0.03)	(0.03)	(0.26)				
1980–1990	0.29	0.17	0.12	0.60				
	(0.01)	(0.04)	(0.04)	(0.13)				
1990–2000	0.10	0.21	−0.10	1.99				
	(0.02)	(0.06)	(0.06)	(0.69)				
1950-2000	0.39	0.35	0.05	0.88				
	(0.02)	(0.05)	(0.05)	(0.13)				

Notes: See the notes for Table 3. In comparison, we here apply an occupation classification scheme based on job titles as opposed to six-digit SOC codes.

B. *Sensitivity Analysis*

In online Appendix E, we consider the sensitivity of the results given in Section IIA to different normalizations and weighting methods, to different subsamples, and to alternative mappings of words to tasks.

In our benchmark decompositions above, we use the frequency of task mentions per 1,000 job ad words as our task measure. While this measure has the advantage of being simple and easy to describe, a potential disadvantage is that different task measures are not directly comparable with one another; the fact that the frequency of nonroutine interactive tasks is 10 times greater than that of nonroutine manual tasks (6.2 nonroutine interactive mentions versus 0.6 nonroutine manual task mentions per 1,000 job ad words) does not necessarily imply that nonroutine interactive tasks are more "important" than nonroutine manual tasks. These differences in magnitudes could instead reflect the breadth of the task word lists. In our first robustness check, we apply a set of normalizations to place task measures on a comparable scale. Specifically, for each job title and year, we normalize its task-related mentions of each individual task by the sum of all mentions across tasks. In other words, we present task content as shares. With these normalizations, as in our benchmark decompositions, we document a substantial shift away from routine tasks and toward nonroutine interactive and analytic tasks. Further, the predominant share of the overall changes in occupational characteristics occurs within rather than between job titles. These results are presented in Tables 19 and 20 in the online Appendix.

In equation (1), we compute our ϑ weights to match the share of workers across four-digit SOC codes. In our second robustness check, our ϑ weights instead reflect job titles' share of vacancies in our dataset. The results are unchanged with this alternate weighting scheme.

Third, throughout our decompositions, we have pooled display ads and classified ads, and we have pooled ads from the *Boston Globe*, *New York Times*, and *Wall Street Journal*. Ads from different regions or in different formats may differ in their task mentions (e.g., display ads tend to mention nonroutine analytic tasks more frequently). Potentially, the changes we report in Section IIA may reflect the changing composition across formats and newspapers. In our third check, we recompute our decompositions separately using two of the main subsamples: *New York Times* classified ads and *New York Times* display ads. While display ads tend to contain a greater frequency of nonroutine analytic and interactive tasks and classified ads contain a greater frequency of nonroutine manual and routine cognitive tasks, there has been a shift toward nonroutine analytic and interactive tasks and away from routine tasks in both sets of ads. Moreover, in both sets of ads, most of the task changes occur within job titles.

Fourth, we recompute Table 4 with Spitz-Oener's (2006) original mapping between tasks and words (i.e., excluding the words we appended from our CBOW model). Again, our conclusions on the shifts in jobs' task contents are unchanged: instead of the 88 percent figure corresponding to Table 4, in this robustness check, 96 percent of the task changes have occurred within job titles.

Fifth, in our online Appendix, we investigate a potential limitation of our approach, namely that we are using job ads (which characterize newly formed jobs) to measure the entire stock of jobs existing at that point in time. Using a perpetual inventory type method, we construct a measure of the stock of each task in each occupation, then recompute the overall and within-occupation shifts in task content. As in our benchmark calculations, the within-job title margin accounts for more than four-fifths of the overall shift in jobs' task content.

C. *Revisiting Autor, Levy, and Murnane* (2003)

We close this section by relating these findings to well-established results in the literature. The decompositions we have performed on our new dataset suggest that a large share of changes in the workplace have taken place within narrowly defined job groupings (either job titles or occupations). A previous literature, however, has established large changes in aggregate task demand coming from shifts in employment shares between occupations.

To show that these two views are consistent, we examine whether our newspaper-based task measures give a portrayal of between-occupation shifts similar to that in the preceding literature. In particular, we replicate figure 1 of Autor, Levy, and Murnane (2003), which reports a key finding in the task literature. In this exercise, industry-gender-education groups are ranked according to the task scores of the occupations in which these groups work.[23] Then, taking the 1960 distribution of employment as the baseline year, Autor, Levy, and Murnane (2003) compute (for each of the five tasks, individually) the employment-weighted mean of the percentiles of the task distribution at different points in time from 1960 to 1998. According to figure 1 of Autor, Levy, and Murnane (2003), nonroutine analytic and interactive task content increases by 8.7 and 12.2 percentiles, respectively, over this period. Aggregate nonroutine manual, routine cognitive, and routine manual task content decreases by 8.7, 5.6, and 0.8 percentiles, respectively. Their figure demonstrates that there has been substantial between-occupation shifts away from routine task-intensive occupations.

In Figure 3, we perform the same exercise, now using our newspaper-based nonroutine and routine task measures. Similar to Autor, Levy, and Murnane (2003), we compute percentiles of demographic groups' task averages based on their 1977 task content. We then compute the mean employment-weighted percentile for each year between 1960 and 2000, taking 1960 employment shares as the baseline. Nonroutine analytic, nonroutine interactive, and routine cognitive task content increase by 4.5, 9.9, and 7.5 percentiles, respectively. Moreover, the aggregate nonroutine manual and routine manual task measures decrease by 13.4 percentiles and 15.8 percentiles, respectively. Overall, the growth rates are similar when using our newspaper data or the DOT: across the five task measures, the correlation between the two sets of growth rates equals 0.55. Pooling across the five task measures and four decades, the correlation in the two sets of decade-by-decade growth rates equals 0.46. The main difference between the data sources is that the estimated change in routine manual tasks in the 1960s and 1970s is −9.2 percentiles in our newspaper data versus 5.6 percentiles according to the DOT.

This exercise indicates that while our decompositions point to within-occupation shifts as an important source of changes in the economy's task content, it is also consistent with one of the foundational results of the task literature—there are substantial between-occupation shifts from routine to nonroutine tasks.

III. Narratives of the Changing Nature of Work

In the previous section, we documented two broad trends. First, a substantial share of the shifts in the tasks that workers perform has occurred within rather than across conventionally defined (six-digit SOC) occupations. Second, using a job title-based categorization—the finest categorization possible—the contribution of

[23] These task scores come from specific questions within the DOT. According to Autor, Levy, and Murnane (2003), GED math scores are a measure of nonroutine analytic tasks; the direction, planning, and control measure corresponds to nonroutine interactive tasks; setting limits, tolerances, and standards is a measure of routine cognitive tasks; finger dexterity is a measure of routine manual tasks; and eye, hand, and foot coordination is a measure of nonroutine manual tasks.

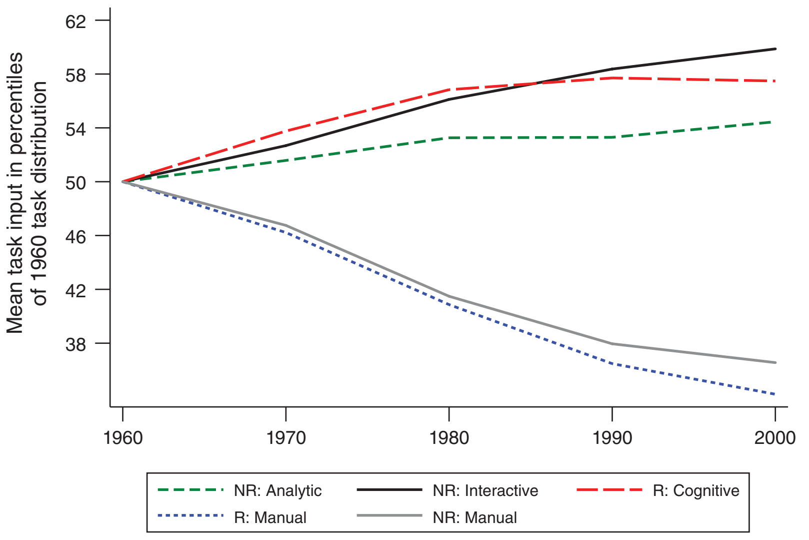

Figure 3. Comparison to Autor, Levy, and Murnane (2003, Figure 1)

Notes: Industry-gender-education groups are ranked by their task content as of 1960. Plotted task percentiles in the succeeding years are employment-weighted averages. There are 2,440 cells representing the combination of two genders, five education groups (<HS, HS, some college, college, postgraduate), and 244 industries (defined by the census ind1990 variable).

within-occupation shifts is equally substantial. Complementing this analysis, in this section, we zoom in, giving concrete examples of the evolution of specific job titles and their associated tasks (Section IIIA). We then discuss the emergence and disappearance of individual job titles (Section IIIB). We intentionally choose a mix of white- and blue-collar jobs to highlight the possible uses of the data and to provide a portrait of the changing nature of work.

A. *Narratives of Task Changes within Jobs*

This section discusses four vignettes, showing that our new data source supports the findings of prior case studies. We depict within-job title task shifts by comparing individual job titles at different points in time.

Our first vignette is motivated by Bartel, Ichniowski, and Shaw's (2007) study of the effect of Computer Numerical Control (CNC) technologies on steel valve manufacturers. According to Bartel, Ichniowski, and Shaw (2007), the introduction of CNC technologies led to a reduction in the demand for worker-performed routine manual tasks. In Panel A of Figure 4, we first plot the frequency of mentions of CNC technologies in machinist job ads.[24] To facilitate comparability across job title characteristics, the plots in this subsection divide each task frequency by the average in our dataset. CNC technologies were rarely, if ever, mentioned up to 1980, then mentioned 0.8 times per 1,000 job ad words in the 1980s and 1.7 times per 1,000

[24] We search for one of the following four strings: "cnc lath*," "cnc mach*," "cnc mill*," or "cnc prog*."

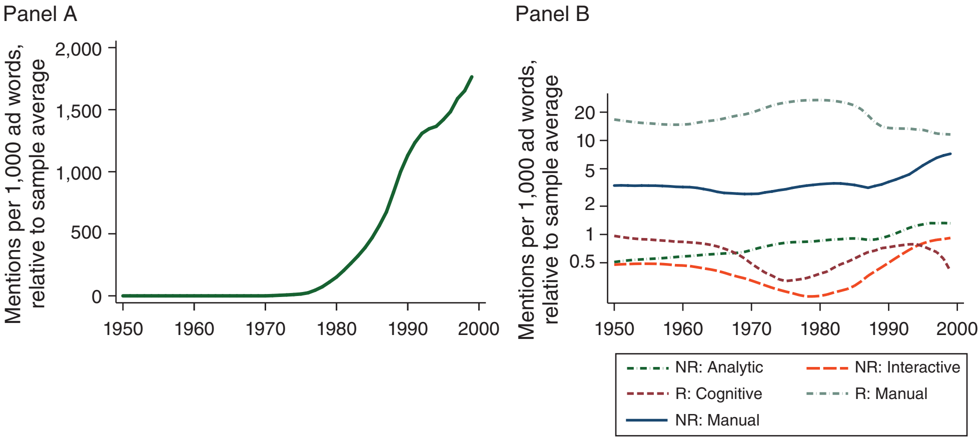

Figure 4. Machinist CNC and Task Measures

Notes: Panel A presents the frequency of mentions of Computer Numerical Control in machinist job ads. Panel B presents the frequency of mentions of Spitz-Oener (2006) tasks among machinist job ads. In both panels, measures are divided by the average frequency, averaging over all job titles and all years. Both panels apply a local polynomial smoother, using a bandwidth of four years. Panel B is plotted on a log scale.

job ad words in the 1990s. These frequencies are 604 and 1,351 times the sample average frequency of CNC technologies. In Panel B, we plot the prevalence of task mentions for machinist workers. Mentions of routine manual tasks were roughly constant for the first few decades of our sample, then fell from 5.5 mentions (20 times the sample average) to 3.4 mentions (12 times the sample average) per 1,000 job ad words between the 1980s and 1990s. On the other hand, nonroutine manual and nonroutine analytic tasks increased in importance beginning in the 1980s. The frequency of nonroutine analytic tasks nearly doubled from 3.8 mentions to 6.6 mentions per 1,000 job ad words between the 1980s and 1990s. In sum, coincident with the diffusion of CNC technologies, machinist jobs shifted away from routine manual tasks toward nonroutine tasks.

Panel A of Figure 5 explores changes in the task content of ads with manager as the job title. Between 1950 and 2000, the frequency of words related to nonroutine interactive tasks in managerial occupations increased modestly. This trend reflects a small increase in the number of words related to selling and a large increase in words the National Research Council (1999) has emphasized in their characterization of the changing nature of managerial work. Summarizing the contemporaneous literature, the National Research Council (1999, 137-38) writes that trends in managerial work involve "the growing importance of skills in dealing with organizations and people external to the firm, … the requirement that [managers] 'coach' … and facilitate relations between workers."

Motivated by this characterization, we plot trends in the mentions of four O*NET work activities in Panel B of Figure 5: working with the public (O*NET Element 4.A.4.a.3), establishing and maintaining relationships (4.A.4.a.4), building teams (4.A.4.b.2), and coaching (4.A.4.b.5). Between the early 1950s and late 1990s, mentions of these four work activities increased by 9 log points (from 9.8 to 10.7

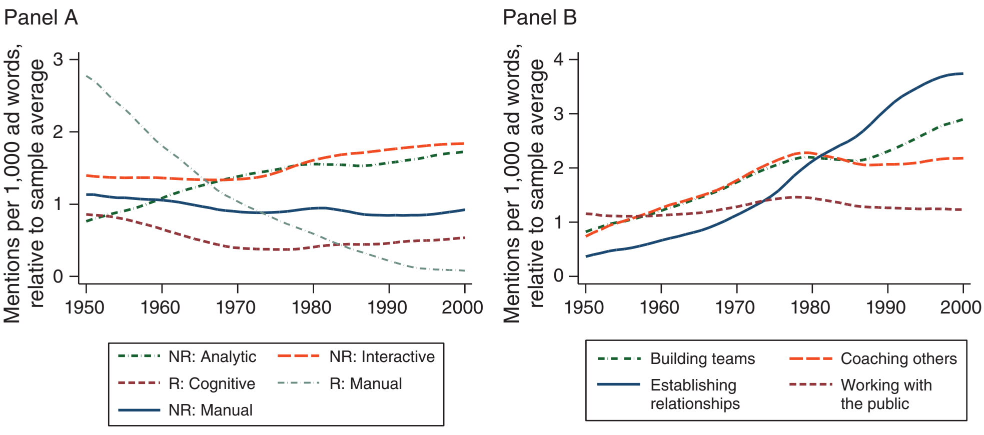

Figure 5. Manager Task Measures

Notes: Panel A presents the frequency of mentions of Spitz-Oener (2006) tasks among ads carrying the manager job title. Panel B presents the frequency of mentions of different O*NET activity measures for manager jobs. In both panels, measures are divided by the average frequency, averaging over all job titles and all years. Both panels apply a local polynomial smoother, using a bandwidth of four years.

mentions per 1,000 ad words of working with the public), 241 log points (from 0.27 to 3.01 mentions per 1,000 ad words of establishing and maintaining relationships), 137 log points (from 1.00 to 3.95 mentions per 1,000 ad words of building teams), and 119 log points (from 0.56 to 1.86 mentions per 1,000 ad words of coaching). In sum, while tasks associated with building and maintaining interpersonal relationships have always been central to managerial occupations, the importance of such tasks has escalated since 1950.

Our third vignette explores changes in the task content of cashier jobs. Over the second half of the twentieth century, organizational and technological shifts altered the environment in which cashiers work. In their study of the retail sector, Basker, Klimek, and Van (2012) chronicle an increase in the prevalence of chains, of general merchandise formats, and of establishment size. These increases in retailer size and scope complemented new technologies—barcode scanners and electronic data interchanges—that reduced the demand for worker-performed routine cognitive tasks. Panel A of Figure 6 confirms this narrative: the frequency of routine cognitive tasks in cashier jobs decreased from 4.3 mentions per 1,000 words in the 1950s (3.4 times the average across all ads and years) to 1.4 mentions per 1,000 words (1.1 times the sample average) in the 1990s. Conversely, the frequency of nonroutine interactive tasks nearly doubled over the sample period.

Our final example discusses a job title for which the task composition is relatively constant: real estate sales. Over the five decades of the sample, the frequency of nonroutine interactive words—the task most central to real estate sales jobs—exhibits no clear trend, increasing from 11.2 mentions per 1,000 job ad words in the 1950s to 13.2 mentions in the 1970s before decreasing to 11.1 mentions in the 1990s. While it is inherently difficult to find prior research that narrates stagnation in an occupation group, we note that Hendel, Nevo, and Ortalo-Magné (2009, 1881) recently

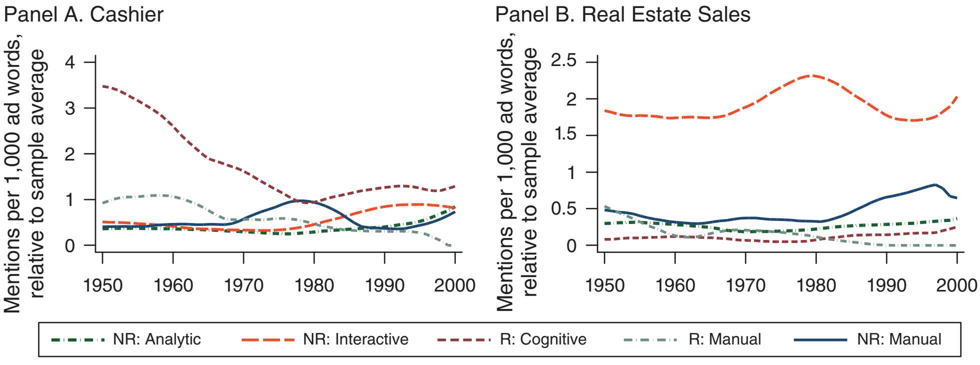

Figure 6. Cashier and Real Estate Sales Task Measures

Notes: In both panels, measures are divided by the average frequency, averaging over all job titles and all years. Both panels apply a local polynomial smoother, using a bandwidth of four years.

described the job of a real estate agent as someone who gives "the homeowner access to a number of services. The National Association of Realtors (NAR) argues that realtors provide valuable help with setting the listing price, preparing the house, checking potential buyers' qualifications, showing the house, bargaining the terms of a deal, and handling the paperwork. Another advantage of working with a realtor is access to the MLS [Multiple Listing Service]." This job description mirrors the 1965 Dictionary of Occupational Titles' definition of a real estate agent as someone who "accompanies prospects to property sites, quotes purchase price, describes features, and discusses conditions of sale or terms of lease. Draws up real estate contracts." Moreover, the Multiple Listing Service mentioned in Hendel, Nevo, and Ortalo-Magné's (2009) quote dates to the late nineteenth century.

To close this section, Table 6 presents a second, complementary illustration of the ways in which job titles have changed over time. We begin by computing the decade-by-decade average task content of each of the job titles discussed in Figures 4 through 6. For instance, we compute that manager jobs in the 1950s contained 4.2 mentions (per 1,000 words) of nonroutine analytic tasks, 8.6 mentions of nonroutine interactive tasks, 1.0 mentions of nonroutine manual tasks, and 0.7 mentions each of routine cognitive tasks and routine manual tasks. For this decade-by-job title combination, we search for the job title for which the task content, averaged over the entire half-century sample period, is closest to this five-dimensional vector.[25] Managerial jobs in the 1950s closely mirrored production manager jobs

[25] The candidate similar job titles are the 200 most frequently mentioned job titles. We apply a Euclidean distance metric. The distance between job title i in decade d and job title j in the 1950–2000 sample is

$$\sum_{h \in \{\text{Spitz-Oener Tasks}\}} \frac{1}{\left(\overline{T}^h\right)^2}\left(\tilde{T}^h_{i,d} - T^h_j\right)^2.$$

In this expression, $\overline{T}^h$ equals the average frequency (in mentions per 1,000 job ad words) of task h across all of the ads in our dataset, $\tilde{T}^h_{i,d}$ equals the frequency of task h in ads in decade d and job title i, and $\tilde{T}^h_j$ equals the frequency of task h in ads in job title j. Dividing by $\left(\overline{T}^h\right)^2$ places all of the tasks on a comparable scale.

Table 6—Near Job Titles

	Frequencies	Similar job title	Frequencies of similar job title
Panel A. Manager			
1950–1959	(4.19, 8.58, 0.70, 1.05, 0.74)	Production manager	(5.43, 8.29, 0.44, 0.36, 0.87)
1960–1969	(5.94, 8.42, 0.61, 0.70, 0.40)	Supervisor	(4.91, 6.63, 0.81, 1.12, 0.37)
1970–1979	(7.08, 8.89, 0.54, 0.44, 0.22)	Manager	(6.95, 10.00, 0.57, 0.64, 0.22)
1980–1989	(7.29, 10.64, 0.57, 0.59, 0.12)	Coordinator	(7.70, 10.16, 0.61, 0.88, 0.05)
1990–2000	(8.10, 11.46, 0.53, 0.64, 0.03)	Coordinator	(7.70, 10.16, 0.61, 0.88, 0.05)
Panel B. Machinist			
1950–1959	(2.64, 3.12, 2.10, 1.14, 4.36)	Machinist	(3.23, 2.52, 2.00, 0.90, 5.29)
1960–1969	(2.87, 2.63, 1.77, 1.07, 4.78)	Machinist	(3.23, 2.52, 2.00, 0.90, 5.29)
1970–1979	(4.00, 1.36, 1.87, 0.40, 7.44)	Machinist	(3.23, 2.52, 2.00, 0.90, 5.29)
1980–1989	(3.78, 1.95, 2.10, 0.57, 5.53)	Machinist	(3.23, 2.52, 2.00, 0.90, 5.29)
1990–2000	(6.61, 5.40, 3.78, 1.22, 3.45)	Electrician	(4.29, 2.63, 4.33, 0.43, 0.76)
Panel C. Cashier			
1950–1959	(1.84, 3.09, 0.25, 4.33, 0.31)	Office assistant	(2.52, 4.23, 0.18, 4.70, 0.29)
1960–1969	(1.55, 2.22, 0.31, 2.26, 0.19)	Cashier	(1.75, 3.14, 0.33, 2.68, 0.20)
1970–1979	(1.16, 1.94, 0.61, 1.50, 0.18)	Computer operator	(2.37, 2.01, 0.62, 1.46, 0.19)
1980–1989	(1.72, 4.13, 0.32, 1.58, 0.07)	Secretary assistant	(2.21, 4.63, 0.17, 1.62, 0.07)
1990–2000	(2.95, 5.58, 0.26, 1.40, 0.10)	Accountant	(2.99, 4.04, 0.27, 1.50, 0.09)
Panel D. Real estate sales			
1950–1959	(1.53, 11.23, 0.28, 0.11, 0.11)	Real estate sales	(1.27, 12.25, 0.28, 0.15, 0.04)
1960–1969	(1.17, 11.14, 0.20, 0.18, 0.05)	Real estate sales	(1.27, 12.25, 0.28, 0.15, 0.04)
1970–1979	(0.90, 13.15, 0.20, 0.06, 0.05)	Real estate sales	(1.27, 12.25, 0.28, 0.15, 0.04)
1980–1989	(1.24, 13.35, 0.20, 0.19, 0.01)	Real estate sales	(1.27, 12.25, 0.28, 0.15, 0.04)
1990–2000	(1.66, 11.09, 0.62, 0.22, 0.00)	Furniture salesperson	(2.65, 9.88, 0.48, 0.31, 0.14)

Notes: The first column gives the five task measures of the job title-decade combination. The five coordinates are: nonroutine analytic, nonroutine interactive, nonroutine manual, routine cognitive, and routine manual. The second column gives the job title—among the 200 most frequently mentioned job titles—that has a task mix (averaged over the whole sample period) that is most similar. The final column gives the task mix for this similar job title.

according to this five-dimensional representation. Averaging over all ads in our sample period, ads with production manager as the job title contained 5.4 mentions of nonroutine analytic tasks, 8.3 mentions of nonroutine interactive tasks, 0.4 mentions of nonroutine manual tasks, 0.4 mentions of routine cognitive tasks, and 0.9 mentions of routine manual tasks. Over time, the nonroutine analytic and interactive task content of "manager" jobs increased. Correspondingly, we find that manager job ads in the 1960s were similar to supervisor ads, while manager ads in the 1970s were similar to the 1950–2000 average of manager ads. By the end of the sample, manager job ads more closely resembled ads for coordinators.

Panels B, C, and D of Table 6 characterize the evolution of machinist, cashier, and real estate sales job ads. Machinist job ads only meaningfully shifted between the 1980s and 1990s. In the last decade of our sample, machinist ads began to resemble 1950–2000 electrician job ads. Cashier job ads from the 1950s and 1960s contained similar task combinations to office assistant and cashier ads. By the 1990s, Cashier job ads more closely resembled ads for accountants.[26]

[26] In online Appendix F, Table 21 presents this same exercise, this time measuring task content as shares. We show that our selected job titles display equally remarkable transitions in this task space, even though the nearest job titles are not identical to those in Table 6.

B. *Emerging and Disappearing Job Titles*

An exceptional feature of our dataset is its ability to characterize the emergence and disappearance of job titles over the second half of the twentieth century.[27] In exploring the evolution of work in the United States, a natural question is whether new job titles differ in their content from older job titles. To the extent that shifts in job title mix occur within conventionally defined occupation codes, existing datasets will understate variation in jobs' task content. In this section, therefore, we explore shifts in the mix of job titles present over our sample period, within six-digit SOCs.

Figure 7 provides four illustrative examples of job title turnover within six-digit SOCs. These examples draw on blue-collar, low-skilled white-collar, and high-skilled white-collar occupations. In Panel A, we plot two job titles within the printing press operator SOC occupation (where the SOC code is 515112). Over the sample period, the share of ads corresponding to the pressman job title declined from 0.23 percent (in the 1950s) to 0.04 percent (in the 1990s). The prevalence of the offset stripper job title increased over the first few decades of the sample, peaking around 1980.[28] Moreover, these job titles were not only placed at different points in time, but also correspond to jobs of different task intensities. The frequency of routine manual tasks is five times higher (0.50 mentions per 1,000 job ad words) in pressman job ads than in offset stripper job ads (0.09 mentions per 1,000 job ad words).

In Panel B, we plot the frequency of data processing and teletype operator job titles, both of which map to the 439021 SOC code. The latter job title comprised 0.05 percent of the job titles in our dataset in the 1950s.[29] Partially as a result of the introduction of the fax machine, low-cost personal computers, and other, newer forms of information and communication technologies, teletype operator jobs essentially disappeared by the 1980s. Within the same six-digit SOC code, the data processing job title emerged in the 1960s. This job title's frequency increased in the 1960s and 1970s, peaked at around 0.09 percent in the early-to-middle 1980s, then declined over the remainder of the sample period. As with job ads corresponding to the printing press operator occupation code, the older vintage teletype operator job ads mention routine tasks more frequently compared to the newer vintage data processing job ads.

Our third set of job titles relates to secretarial and administrative assistant occupations. With the introduction of word processing equipment and software, job titles specifically relating to typing have declined in frequency over the sample period. In their place, job titles denoting interaction with visitors or clients have increased

[27] Lin (2011) compares vintages of the DOT and the US Census Classified Indexes to classify new job titles as they appear over long horizons. Relative to Lin (2011), we are able to characterize the task and skill content of new job titles. Beyond identifying emerging job titles, we are also able to identify disappearing jobs, both of them at higher frequencies than Lin's approach allows.

[28] Version 4.0 of O*NET contains a separate eight-digit SOC code, 515022.06, associated with strippers. According to this version of O*NET, strippers are workers who "Cut and arrange film into flats (layout sheets resembling a film negative of text in its final form) which are used to make plates. Prepare separate flats for each color." The introduction of digital prepress technologies has largely replaced the tasks performed by offset strippers.

[29] A teletype machine is "a printing device resembling a typewriter that is used to send and receive telephonic signals—formerly a US registered trademark." Merriam-Webster Dictionary (2019).

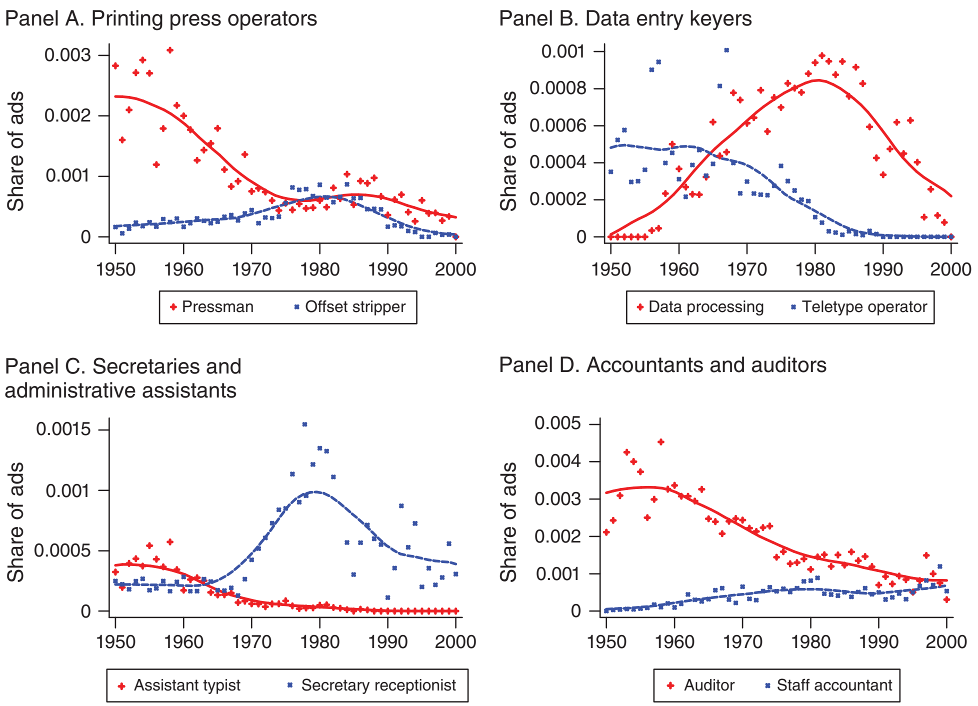

Figure 7. Job Titles' Vacancy Shares

Notes: Within each panel, we plot the raw share and smoothed frequency of two job titles within the same six-digit SOC. In the four panels, we plot job titles from printing press operators (SOC 515112), data entry keyers (SOC 439021), secretaries and administrative assistants (SOC 436014), and accountants and auditors (SOC 132011). Within each panel, we apply a local polynomial smoother with a bandwidth of four years.

in frequency. While both assistant typist and secretary receptionist jobs are heavily centered on routine cognitive tasks, secretary receptionist jobs contain a greater frequency of nonroutine interactive tasks and fewer routine cognitive tasks.

Our final example compares two occupations that fall within the accountants and auditors SOC occupation code: auditor and staff accountant. Between the 1950s and 1990s, the share of ads referring to staff accountant positions increased sixfold, from 0.01 percent to 0.06 percent. Over the same period, the share of auditor ads fell from 0.33 percent to 0.09 percent. While both staff accountants and auditors work with firms' financial statements, auditors' work centers more on verification rather than preparation. Moreover, these differences are reflected in our Spitz-Oener-based task measures. Averaging over the ads within our newspaper text, auditor positions include 3.3 mentions of nonroutine analytic tasks per 1,000 job ad words, which is half of that for staff accountant jobs.

The key takeaway from this analysis is that even within six-digit SOC codes, there has been a substantial transformation in the composition of ads' job titles. This transformation holds more broadly: of the job titles that were the most common within their six-digit SOC in the 1950s, nearly 40 percent had completely disappeared by the 1990s.[30] Similarly, of the job titles that were the most common within their six-digit SOC in the 1990s, more than 45 percent were not in existence in the 1950s.[31]

A second takeaway from these vignettes is that compositional changes in job titles are not merely cosmetic but instead represent real changes in occupational tasks. In particular, ads corresponding to newer vintage job titles contain a greater share of nonroutine tasks and a lesser share of routine tasks.

We next explore the extent to which this second takeaway is a more systematic feature of occupational change. We first define v_j^p, *vintages* of job title j, as the pth quantile of the distribution of years in which the job title appears in our data; in computing these quantiles for each job title, we weight according to the job title's share of ads $(S_{j,t})$ in each year. For p close to 0, v_j^p compares different job titles based on when they first emerged in our dataset. In contrast, v_j^p for p close to 1 compares job titles based on their disappearance from our dataset. For instance, while secretary receptionist and offset stripper were both increasing in frequency between the 1950s and 1980s, the offset stripper job title disappeared before secretary receptionist jobs did. Correspondingly, $v_j^{0.05} = 1955$ for both job titles. But $v_j^{0.95} = 1989$ for offset stripper jobs, which is lower than the value for secretary receptionist jobs, 1997. On the other hand, since offset stripper jobs both emerged and disappeared before staff accountant jobs, v_j^p is greater for $j =$ staff accountant than for $j =$ offset stripper for both $p = 0.05$ and $p = 0.95$.

With these definitions in hand, we regress our task measures against our job title vintage measures:

$$task_j^h = \beta_o + \beta_1 v_j^p + \varepsilon_{j,h}. \tag{2}$$

In equation (2), $task_j^h$ measures the average number of mentions of task h per 1,000 job ad words in job title j's ads over the sample period, and β_o are SOC fixed effects. Table 7 reports the results of this regression. New job titles are associated with a greater frequency of nonroutine analytic and interactive tasks and a lower frequency of routine cognitive and routine manual tasks.[32] These patterns hold both within and across occupations (panels A, C, and E versus panels B, D, and F) and for our three job title vintage measures (panels A and B versus panels C and D versus

[30] Linotype operator, nurse governess, and office boy were the most common job titles within their six-digit SOCs—515111, 311011, and 35021, respectively—in the 1950s. None of these job titles was present in any job ads in the 1990s. Among the set of disappearing job titles, these three were the most common job titles in the 1950s (as a share of all 1950s job ads).

[31] Nurse practitioners, medical billers, and telemarketers were the most common job titles within their SOC codes—291171, 292071, and 419041, respectively—in the 1990s. There were no ads corresponding to these job titles in the 1950s.

[32] Using the coefficient estimates from panel D, a one-decade increase in the entry vintage is associated with a 0.12 standard deviation reduction in mentions (per 1,000 job ad words) of routine cognitive tasks, a 0.28 standard deviation reduction in the frequency of routine manual tasks, and a 0.17 standard deviation increase in mentions of Deming and Kahn's computer skills.

Table 7—Relationship between Task Measures and Job-Title Vintages

	Nonroutine			Routine		Deming and Kahn
Dependent variable	Analytic	Interactive	Manual	Cognitive	Manual	Computer
Panel A. No fixed effects, $p = 0.05$						
Coefficient	0.032	0.046	0.001	−0.029	−0.008	0.076
Standard error	0.007	0.008	0.001	0.006	0.001	0.003
Panel B. Six-digit SOC fixed effects, $p = 0.05$						
Coefficient	0.011	0.024	−0.001	−0.015	−0.007	0.054
Standard error	0.002	0.002	0.000	0.001	0.000	0.001
Panel C. No fixed effects, $p = 0.50$						
Coefficient	0.050	0.081	0.001	−0.041	−0.013	0.101
Standard error	0.007	0.008	0.001	0.005	0.001	0.003
Panel D. Six-digit SOC fixed effects, $p = 0.50$						
Coefficient	0.036	0.049	−0.001	−0.020	−0.010	0.090
Standard error	0.001	0.002	0.000	0.001	0.000	0.001
Panel E. No fixed effects, $p = 0.95$						
Coefficient	0.027	0.047	−0.001	−0.013	−0.008	0.058
Standard error	0.006	0.007	0.001	0.005	0.000	0.004
Panel F. Six-digit SOC fixed effects, $p = 0.95$						
Coefficient	0.025	0.028	0.000	−0.003	−0.005	0.052
Standard error	0.001	0.002	0.000	0.002	0.000	0.001

Note: Within each panel and column, we present coefficient estimates and standard errors corresponding to estimates of equation (2).

panels E and F). In the final column of Table 7, we demonstrate that ads with newer vintage job titles contain a greater frequency of computer-related words.[33]

The changing composition of job titles also illustrates the different phases of the digital revolution. In Figure 8, we plot four job titles that refer to different aspects of the development of information and communication technologies. Job ads for software engineers first appeared around 1970, became increasingly common in the 1970s, and had no clear increase thereafter. Appearing in the 1970s and 1980s, developer and database administrator jobs grew rapidly in the 1990s. Finally, mirroring the diffusion of network technologies in the 1980s and 1990s, network engineer positions only emerged in the late 1980s.[34] The rapid diffusion of new job titles in computer occupations accords with Deming and Noray's (2018) characterization of 1983–1992 as a period of transformation within STEM occupations.[35]

[33] Online Appendix F replicates this exercise with job title-year pairs as the unit of observation and with year fixed effects included in our regression specification. The directional results are for the most part unchanged, though with smaller magnitudes than in Table 7. A reason for these differences is that the specification we explore in online Appendix F requires that the job titles being compared be observed within the same year, thus removing all information coming from nonoverlapping job titles.

[34] Unplotted, and appearing even more recently than network engineer jobs, are job titles specifically referring to the internet. A majority of the ads corresponding to web developer, web designer, and web master jobs were placed after 1998.

[35] Deming and Noray (2018) use our dataset to construct measures of change within occupations over time. STEM occupations include not only computer-related occupations but also occupations within SOC codes

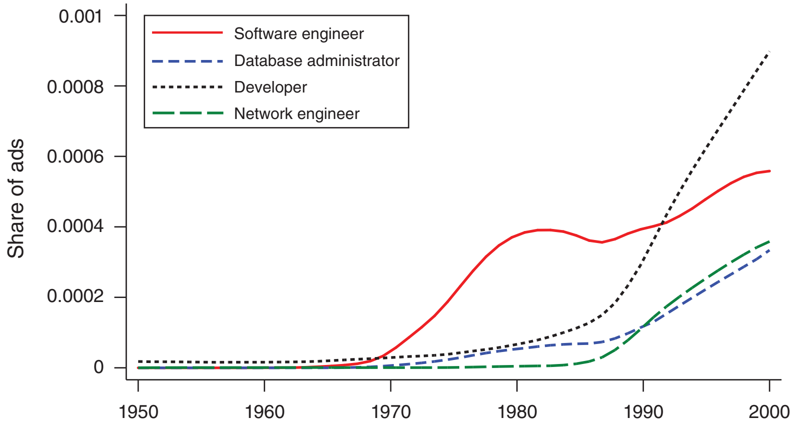

FIGURE 8. COMPUTER JOB TITLES' VACANCY SHARES

Notes: We plot the share of ads corresponding to each job title. We apply a local polynomial smoother with a bandwidth of four years.

IV. Conclusion

In this paper, we introduce a new dataset and use it to chronicle changes in US occupations and job tasks between 1950 and 2000. We document that a predominant share of changes in the task composition of the workforce has occurred within rather than between occupations. Beyond the decline that occupations intensive in routine tasks have experienced as a share of the workforce—a central pattern of the existing task literature—individual occupations' routine task content has declined as well. We also show that while our findings resonate with previous findings for individual case studies, our new dataset readily lends itself to a more exhaustive analysis of the evolution of the labor market than is possible based on standard data sources.

Beyond this project, our newspaper data have the potential to address other economic questions related to the labor market. For example, in related work, we use the text to measure the adoption of new computer technologies in order to study how these technologies interact with the task content of jobs (Atalay et al. 2018). More generally, we view our newspaper-based job vacancy text as offering an opportunity to study, over a longer time horizon, questions that have been examined using online job vacancies.

APPENDIX A: DATA USERS' GUIDE

This is a guide to the dataset introduced in "The Evolution of Work in the United States."

beginning with 17 or 19. Moreover, their measures of occupational change will encapsulate not only within-job-title changes in skills and tasks but also shifts in job title composition similar to those depicted in Figure 8.

A. *Site and Contents*

The site https://occupationdata.github.io contains all materials related to the dataset:

- Details on the procedure by which we (i) process the digitized text, (ii) classify ads as job ads versus other types of advertisements, (iii) determine the boundaries of each individual ad, (iv) identify the job title within each job ad, (v) map words to job characteristics, and (vi) map job titles to occupational codes.
- Python notebooks that implement the six-step procedure.
- The file apst_mapping.xls, which details the mapping between raw text and job characteristics.
- The final analysis datasets.

B. *Dataset Vintages*

The site currently hosts the third version of our dataset, uploaded on May 15, 2019. The site also archives the two previous versions we have released: the first version, which was made available on July 1, 2017, and the second version, which was made available on April 7, 2018. While we do not anticipate major changes in the future, we may add additional variables over time. We will continue to archive older versions of our dataset.

C. *Dataset Contents*

The data page contains eight downloadable datasets in which the data are aggregated to the following levels: (i) job title, (ii) job title × year pair, (iii) SOC code occupation, (iv) SOC code × year pair, (v) OCC code occupation, (vi) OCC code × year pair, (vii) job title × source pair, and (viii) job title × year × source triple.[36] While the sample period in the published article is 1950 to 2000, the datasets contain information from ads that date to 1940.

The list to follow explains the variables in our dataset. For each variable that measures job characteristics, we provide the number of word mentions per ad that correspond to a particular job characteristic. Each of the eight downloadable data datasets contains the following sets of variables:

- activity_*, requirement_*, skill_*, style_*: These variables correspond to different O*NET Work Elements. Variables that end with a *_C in their name include words taken from our continuous bag of words model.
- big5_*: These variables correspond to "Big 5" traits. We use the categorization of words to Big 5 traits described by John, Naumann, and Soto (2008).

[36] Users of our dataset should be able to reproduce the third, fourth, fifth, or sixth datasets from the first two datasets. To do so, users can merge the job-title-based datasets with the mapping between job titles, SOC codes, and OCC codes that we provide. Here, "source" refers to whether the job ad appears as a *Boston Globe* classified ad, a *Boston Globe* display ad, a *New York Times* classified ad, a *New York Times* display ad, or a *Wall Street Journal* classified ad. Users of our dataset should be able to reproduce the first or second datasets from the last two datasets.

- deming_*: These variables correspond to the skill measures discussed in Deming and Kahn (2018). Variables that end with a *_C in their name include words taken from our continuous bag of words model.
- degree_*: These variables correspond to seven potential degrees: associate's, bachelor of arts, bachelor of science, master's, MBA, PhD, and CPA.
- experience_*: These variables correspond to experience requirements—whether employers ask for 1 year, 2 years, 3 years, 4 years, or 5+ years of experience.
- oth_* : These variables correspond to various job characteristics, including: the hours of the job (both the total number and the actual schedule), whether the applicant is asked for a salary history, and whether the employer offers tuition reimbursement.
- spitz_*: These variables correspond to nonroutine (analytic, interactive, or manual) or routine (cognitive or manual) tasks. Variables that end with a *_C in their name include words taken from our continuous bag of words model.
- technology_*: These variables count the number of mentions of different pieces of technology. The 48 technologies are listed in Atalay et al. (2018).
- length; words: The first of these variables counts the number of correctly spelled words, incorrectly spelled words, and nonword tokens per ad. The second of these variables counts the number of correctly spelled words per ad. In work that uses our dataset, we suggest normalizing per-ad job measures by the number of correctly spelled words per ad.
- ct2: This variable gives the number of ads corresponding to the given observation.

REFERENCES

Abraham, Katharine G., and Michael Wachter. 1987. "Help-Wanted Advertising, Job Vacancies, and Unemployment." *Brookings Papers on Economic Activity* 17 (1): 207–48.

Acemoglu, Daron, and David Autor. 2011. "Chapter 12—Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1043–1171. Amsterdam: North-Holland.

Anastasopoulos, Jason, George J. Borjas, Gavin G. Cook, and Michael Lachanski. 2018. "Job Vacancies, the Beveridge Curve, and Supply Shocks: The Frequency and Content of Help-Wanted Ads in Pre- and Post-Mariel Miami." NBER Working Paper 24580.

Atalay, Enghin, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. 2018. "New Technologies and the Labor Market." *Journal of Monetary Economics* 97: 48–67.

Autor, David H. 2013. "The 'Task Approach' to Labor Markets: An Overview." *Journal for Labour Market Research* 46 (3): 185–99.

Autor, David H. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29 (3): 3–30.

Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *American Economic Review* 96 (2): 189–94.

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

Bartel, Ann, Casey Ichniowski, and Kathryn Shaw. 2007. "How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills." *Quarterly Journal of Economics* 122 (4): 1721–58.

Basker, Emek, Shawn D. Klimek, and Pham Hoang Van. 2012. “Supersize It: The Growth of Retail Chains and the Rise of the ‘Big-Box’ Store.” *Journal of Economics and Management Strategy* 21 (3): 541–82.

Brynjolfsson, Erik, and Andrew McAfee. 2014. *The Second Machine Age—Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: W.W. Norton and Company.

Cortes, Guido Matias, Nir Jaimovich, and Henry E. Siu. 2018. “The ‘End of Men’ and Rise of Women in the High-Skilled Labor Market.” NBER Working Paper 24274.

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

Deming, David, and Lisa B. Kahn. 2018. “Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals.” *Journal of Labor Economics* 36 (S1): S337–69.

Deming, David J., and Kadeem L. Noray. 2018. “STEM Careers and the Changing Skill Requirements of Work.” NBER Working Paper 25065.

DeVaro, Jed, and Oliver Gürtler. 2018. “Advertising and Labor Market Matching: A Tour through the Times.” *Journal of Labor Economics* 36 (1): 253–307.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2014. “Occupational Tasks and Changes in the Wage Structure.” IZA Discussion Paper 5542.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. “Text as Data.” *Journal of Economic Literature* 57 (3): 535–74.

Hendel, Igal, Aviv Nevo, and François Ortalo-Magné. 2009. “The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com.” *American Economic Review* 99 (5): 1878–98.

Hershbein, Brad, and Lisa B. Kahn. 2018. “Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings.” *American Economic Review* 108 (7): 1737–72.

John, Oliver P., Laura P. Naumann, and Christopher J. Soto. 2008. “Paradigm Shift to the Integrative Big Five Trait Taxonomy: History, Measurement, and Conceptual Issues.” In *Handbook of Personality: Theory and Research*, 3rd ed., edited by Oliver P. John, Richard W. Robins, and Lawrence A. Pervin, 114–58. New York: Guilford Press.

Lin, Jeffrey. 2011. “Technological Adaptation, Cities, and New Work.” *Review of Economics and Statistics* 93 (2): 554–74.

Marinescu, Ioana, and Ronald Wolthoff. Forthcoming. “Opening the Black Box of the Matching Function: The Power of Words.” *Journal of Labor Economics*.

Merriam-Webster Dictionary. 2019. Teletype. Merriam-Webster, Inc. https://www.merriam-webster.com/dictionary/teletype.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding. 2018. “Task Specialization in U.S. Cities from 1880 to 2000.” *Journal of the European Economic Association* 17 (3): 754–98.

Miller, Anne R., Donald J. Treiman, Pamela S. Cain, and Patricia A. Roos, eds. 1980. *Work, Jobs, and Occupations—A Critical Review of the Dictionary of Occupational Titles*. Washington, DC: National Academies Press.

Modestino, Alicia S., Daniel Shoag, and Joshua Ballance. Forthcoming. “Upskilling: Do Employers Demand Greater Skill When Skilled Workers Are Plentiful?” *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_00835.

National Research Council. 1999. *The Changing Nature of Work—Implications for Occupational Analysis*. Washington, DC: National Academies Press.

National Research Council. 2010. *A Database for a Changing Economy—Review of the Occupational Information Network (O*NET)*. Washington, DC: National Academies Press.

Ross, Matthew B. 2017. “Routine-Biased Technical Change: Panel Evidence of Task Orientation and Wage Effects.” *Labour Economics* 48: 198–214.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. “Integrated Public Use Microdata Series: Version 6.0 [dataset].” IPUMS, Minneapolis, Historical Census Projects, University of Minnesota. http://doi.org/10.18128/D010.V6.0 (accessed January 18, 2018).

Spitz-Oener, Alexandra. 2006. “Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure.” *Journal of Labor Economics* 24 (2): 235–70.

This article has been cited by:

1. J. Carter Braxton, Bledi Taska. 2023. Technological Change and the Consequences of Job Loss. *American Economic Review* **113**:2, 279-316. [Abstract] [View PDF article] [PDF with links]
2. Sara B. Heller, Judd B. Kessler. 2022. Soft Skills in the Youth Labor Market. *AEA Papers and Proceedings* **112**, 121-125. [Abstract] [View PDF article] [PDF with links]