The Evolving U.S. Occupational Structure

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

September 25, 2017

Abstract

Using the text from help wanted ads, we construct a new data set of occupational task content from 1960 to 2000. We document that within-occupation task content shifts are at least as important as employment shifts across occupations in accounting for the aggregate decline of routine tasks. Motivated by these patterns, we first apply our new task measures to a reduced-form statistical decomposition. We then embed our measures in an equilibrium model of occupational choice. These two exercises indicate that shifts in the relative demand for tasks account for a 25 log point increase in 90-10 earnings inequality over our sample period. **JEL Codes: J24, M51, O33**

1 Introduction

Labor income inequality in the United States has increased considerably over the last several decades. Between 1960 and 2000, the ratio of earnings for the median worker, compared to the worker at the 10th percentile of the earnings distribution, has increased from 2.75 to 2.86. Over the same period, the 90-50 earnings ratio increased even more starkly, from 1.81 to 2.27. Alongside rising inequality, there have been dramatic changes in the composition of U.S. employment. For example, the employment share of occupations intensive in routine tasks has shrunk, while the share of occupations emphasizing nonroutine tasks, as well as social and cognitive skills, has grown (Autor, Levy, and Murnane, 2003; Autor, Katz, and Kearney, 2005; and Autor and Dorn, 2013).

*Atalay and Phongthiengtham: Department of Economics, University of Wisconsin-Madison. Sotelo: Department of Economics, University of Michigan-Ann Arbor. Tannenbaum: Department of Economics, University of Nebraska-Lincoln. We thank Sara Heller, Andrei Levchenko, Pablo Ottonello, and Chris Taber for constructive and helpful comments, and Erin Robertson for excellent editorial assistance. We also thank seminar audiences at Geneva, Lima, Michigan, Penn State, and Wisconsin. We acknowledge financial support from the Washington Center for Equitable Growth.

The evidence thus far points to the decline of employment in U.S. occupations intensive in routine tasks, but is largely silent on whether the content of occupations themselves has changed. Meanwhile, case studies of individual occupations have found considerable changes. For example, a report by the National Research Council (1999) suggests that managerial occupations over the second half of the twentieth century increasingly have emphasized team management, coaching skills and tasks, and interaction with customers. The rise in these skills and tasks were accompanied by a decline of managerial tasks related to direct control of subordinates.[1] These findings raise the question of whether managerial occupations are unique in experiencing changes in tasks, or if comparable changes have occurred elsewhere. Moreover, to the extent that the tasks that workers perform on the job have shifted, what are the implications for the distribution of earnings?

In this paper we show that, in fact, substantial changes in task composition did occur within occupations since 1960.[2] Moreover, we find that these within-occupation changes in task content account for much of the observed increase in earnings inequality. We start by constructing a new data set using the text content of approximately 4.1 million job ads appearing in three major metropolitan newspapers — the *New York Times*, the *Wall Street Journal*, and the *Boston Globe*. We then map the words contained in job ad text to different classifications of task content. Our main strategy relies on Spitz-Oener's (2006) classification of words into routine (cognitive and manual) and nonroutine (analytic, interactive, and manual) tasks. We complement this approach using the classifications in Deming and Kahn (2017), Firpo, Fortin, and Lemieux (2014), and the Occupational Information Network (O*NET). This last mapping is useful, since O*NET is at present the de facto standard in the literature for measuring the content of occupations. To validate our new data set, we demonstrate that the measures that result from our mapping correlate, across occupations, with O*NET's measures of the importance of different tasks.[3]

[1] The National Research Council report chronicles the evolution of several other occupations. In Section 4, we link the changes within managerial, clerical, and assembly occupations documented in the report to the new measures we introduce in this paper.

[2] 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." (p. 1045, emphasis in the original) We adopt these definitions of skills and tasks throughout our paper.

[3] We also perform several checks on the data to argue that neither the selection of ads into newspapers nor the fact that our data originate in metropolitan areas bias our results. First, we show that the distribution of ads across occupations is similar to the distribution of employment according to the decennial census (and our results largely agree with analogous comparisons made with online vacancies, e.g. Hershbein and Kahn, 2016). Second, we show that the educational requirements by occupation in our data set correlate reasonably well with those observed in the census (Appendix A). Third, we show that trends in the propensity of unemployed workers to search for jobs through help wanted ads does not vary with the task content of their prior occupation (Appendix B). And, fourth, we show that our main results are not driven by the fact

We show that, in our database of newspaper ads, words related to interactive tasks have been increasing in frequency, while words related to routine tasks (especially routine manual tasks) have declined in frequency between 1960 and 2000. For instance, we find that the frequency of words related to routine cognitive tasks has declined by almost half from the early 1960s to the late 1990s, from 1.00 mentions per thousand job ad words to 0.54 mentions per thousand words. The frequency of routine manual tasks has declined even more starkly. On the other hand, the frequency of words related to nonroutine interactive tasks has increased by more than 54 percent, from 3.19 to 5.48 mentions per thousand of job ad words. Importantly, we find that a large fraction of the aggregate changes in both nonroutine and routine task related words has occurred within occupations, rather than through changes in the occupations' employment shares.

Having documented these patterns, we provide two alternative quantifications of the relation between changes in the task content of occupations and the rise of inequality in the U.S. The first one relies on decomposition methods applied to the earnings distribution; the second, on an equilibrium model of worker occupational choices and comparative advantage.

In our first quantification, we incorporate our occupational measures into Fortin, Lemieux, and Firpo (2011)'s methodology for decomposing changes in the wage distribution across points in time. Using these methods, we break down changes in the distribution of earnings over time into changes in the attributes of workers and their occupations (the "composition" effect), as well as changes in the implicit prices of those observable characteristics (the "wage structure" effect). Next, we further break down changes in the distribution of earnings into the contributions of observable characteristics (e.g., the contribution of changes in educational attainment and of changes in the returns to education).

Our results suggest that, relative to the upper tail of the income distribution, tasks which were valued highly in low-wage jobs (in particular manual tasks) declined in importance. These task changes, however, come more from changes within the occupations themselves than from shifts of employment across occupations. In the decomposition, changes in occupational task content account, through composition effects, for a 32 log point increase in 90-10 male earnings inequality. Wage structure effects — changes in tasks' implicit prices — account for a 6 percentage point decrease in 90-10 inequality. Among the 32 percentage points that are due to composition effects, routine manual and nonroutine interactive tasks respectively account for a 24 and 8 log point increase in inequality. (The contribution of changes in routine cognitive, nonroutine analytic, and nonroutine manual tasks is more modest.) If we instead ignore the evolution of occupational characteristics across time, and keep the task measures fixed, we can only account for a total of 3 percentage points of the

that most of our ads come from a selected number of large metro areas (Appendix E.3).

observed increase in 90-10 inequality. Hence occupational measures of task intensity account for a large component of the increase in earnings inequality between 1960 and 2000, but only if task measures are allowed to vary across time within occupations.

In our second quantification, which we view as complementary to the earnings distribution decomposition, we construct a general equilibrium model of occupational choice. In our model, individual occupations are represented as a bundle of tasks. Workers' skills govern their abilities to perform each of the individual tasks in their occupation, and give rise to comparative advantage. These skill levels are functions of workers' observable characteristics — like gender, education, and experience — but also contain an idiosyncratic component. Based on their skill levels and the demand for tasks within each occupation, workers select into the occupation with the highest payoff.

We estimate the skills by which each demographic group performs different groups of tasks, combining information on demographic groups' wages and occupation choices. Using our estimated model, we calculate that changes in the relative demands for tasks have led to a 25 log point increase in inequality among men, and a 22 log point increase in inequality among women and men. The intuition is that less educated workers both tend to be at the bottom of the wage distribution and have a comparative advantage in manual tasks, while the demand for the latter has declined.

Both the statistical decompositions and our model-based counterfactual exercises have their own advantages. In spite of their differences, however, both methods indicate that a relative decline in the demand for routine tasks has substantially increased earnings inequality between 1960 and 2000. Moreover, both of these exercises require information on changes in occupations' task content, something that our data set is suited to measure.

Our paper builds on two literatures, the first of which 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-skill workers. More recently, Firpo, Fortin, and Lemieux (2014) decompose changes in the distribution of wages into the contribution of occupational characteristics and other factors (including de-unionization, changes in minimum wage, and changes in worker demographics). Deming (2017) documents that employment and wage growth has been confined to occupations which are intensive in both social and cognitive skills. Michaels, Rauch, and Redding (2016) extend Autor, Levy, and Murnane (2003) to study changes in employment shares by task content over a longer time horizon. They use a methodology that is related to ours, using the verbs from the Dictionary of Occupational Titles' occupational descriptions and their thesaurus-based meanings. Burstein, Morales, and Vogel (2015) quantify the impact of the adoption of computers on between-group wage inequality.

Because commonly available data sets measuring occupational characteristics are infrequently and irregularly updated, in most of these papers occupational characteristics are fixed over time.[4] One paper that focuses on within-occupation changes is Spitz-Oener (2006), which uses survey data from four waves of German workers to track skill changes within and between occupations, from the late 1970s to the late 1990s. One of our contributions is to study these changes in the U.S. context, extending the Spitz-Oener (2006) task classifications, and examining the implications of the observed changes in tasks for the wage distribution. Our paper contributes more generally to the measurement of time-varying task content of occupations by using newspaper job ads, a rich and largely untapped source of data.[5] Newspaper data have the advantage of appearing in the field: Firms post these ads while they are actively searching for workers, meaning that, unlike survey data, our data are not subject to recall-bias or selective nonresponse. Our new data set contains rich information on U.S. occupations' skill requirements, technology usage, work activities, work styles, and other job characteristics, between 1960 and 2000. These data are complementary to existing data sources currently used to study the evolution of the U.S. labor market.[6]

A second related literature 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.[7] Using data from CareerBuilder, Marinescu and Wolthoff (2016) document substantial variation in job ads' skill requirements and stated salaries within narrowly-defined occupation codes. Moreover, the within-occupation variation is to a large extent explained by the vacancy posting's job title. Using online job ads, Hershbein and Kahn (2016) and Modestino, Shoag, and Ballance (2016) argue that jobs' skill requirements have increased during the post-Great Recession period; Deming and Kahn (2017) find that firms that post

[4] Ross (2017) applies the 2003 to 2014 vintages of the O*NET database to construct a panel of occupations' skill and task intensity measures. In Ross (2017), within-occupation differences in task intensities are applied as an alternative source of variation, relative to previously used between-occupation trends, in identifying declines in the price of workers' routine task engagement in the last decade. In earlier work, Autor, Levy, and Murnane (2003) use the 1977 and 1991 versions of the Dictionary of Occupational Titles to compare changes in occupations' task content and computer adoption rates within these occupations.

[5] Also applying the text from newspaper ads, DeVaro and Gurtler (2017) document that before 1940 both job seekers and firms posted advertisements to match with one another. Only since 1940 has it been the case that firms were the primary party posting ads.

[6] Our new data set can be found at http://ssc.wisc.edu/~eatalay/occupation_data .

[7] Text analysis has been fruitfully applied in other branches of economics. Gentzkow and Shapiro (2010) study the text from newspapers and from the Congressional Record to construct an index of newspapers' partisan slant. Hoberg and Phillips (2016) use the text from publicly-traded firms' filings to the Securities and Exchange Commission to classify industries and to draw insights about the identities of firms' competitors.

ads with a high frequency of words related to social and cognitive skills have higher labor productivity and pay higher wages. By bringing in newspaper data, we contribute to this literature with measurement of long-run changes in tasks and skills in the labor market.

The rest of the paper is organized as follows. Section 2 outlines the construction of our data set of occupations and their skill and task content. We compare this new data set to existing data sources in Section 3. In Section 4, we document changes in the distributions of occupational skills and tasks. Section 5 links task changes to changes in the earnings distribution, first in a reduced-form exercise and then with the aid of a structural model. Section 6 reviews our results and suggests areas for future research.

2 A New Data Set of Occupational Characteristics

In this section, we discuss the construction of our structured database of occupational characteristics. The primary data sets are raw text files, purchased from ProQuest and originally published in the *New York Times*, *Wall Street Journal*, and *Boston Globe*. We complement these text files with a data set purchased from Economic Modeling Specialists International (EMSI) for the purpose of identifying each ad's occupation code. We extract from the newspaper text information about the frequency of words related to occupations' skill requirements and their work activities. The first step in our approach is to clean and process the raw newspaper data. Then, we map job ad titles to Standard Occupational Classification (SOC) codes. Finally, we map job ad text into tasks, by grouping sets of words according to their meaning. Once we describe these procedures, we present some simple descriptive statistics from our constructed data set.

2.1 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 newspaper into text files. Figure 1 presents a snippet of the raw text from a page of display ads from the 1979 *Boston Globe*. This text refers to three vacancy postings, one for a "Mutual Fund Clerk" position, a second for a "Payments Clerk," and a third for a "Terminal Operator." 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), there is still quite a lot of information to be extracted. Common work activities such as typing, processing applications, or maintaining client records are mentioned. In addition, one of the three advertisements includes an experience requirement. The raw text files that ProQuest has provided us allow us to

Figure 1: Unprocessed Text from the *Boston Globe*, November 4, 1979, Display Ad #133

```
rapid growth through continued innovation and diversification If you
are highly motivated person who takes pride in your work and the company
you work for consider career with Fidelity \n MUTUAL FUNDS CLERKS \n
Individuals with 1-2 years funds transfer experience to process Keogh
IRA accounts and adjustments \n PAYMENTS CLERKS \n Varied les processing
new account applications and payments \n StI nt flt ner fr nlr ct tfl \n
Mnirk \n and maintaining client \n strong record-keeping 50 wpm \n Data
Control Department \n sorting and \n Brokerage \n environments with \n
benefits package for our Boston convenient to the Market \n gr WilnliE
lU5lty \n success \n -l 1?Q1.a1 ol P-1Sl1v MPloV \n Fidelit \n Group \n
82 DEVONSHIRE STREET BOSTON MA 02109 \n 111 \n -l \n TERMINAL OPERATOR
\n Position involves typing policy related information Into computer
terminal No previous computer experience required Typing 5055 wpm
Excellent benefits plus work Incentive program in addition to starting
salary of S150-165.
```

Notes: The figure presents text from the first three vacancy postings in a page of display ads in the *Boston Globe*. An "\n" refers to a line break.

isolate the subset of text that come from advertisements, but do not directly allow us to identify which ads are job ads (as opposed to other types of advertisements). Nor does the text indicate when one job ad ends and another one 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.

Distinguishing vacancy postings from other advertisements. To distinguish vacancy postings from other types of advertisements, we apply a Latent Dirichlet Allocation (LDA) model. An LDA model classifies documents to one of a number of topics. In these models, the probability that different words appear in a document depends on the hidden topic of that document. LDA model estimation yields estimates of these conditional probabilities.

In our context, the LDA model is used to distinguish pages of job ads (one of the model's topics) from other groups of ads. Estimation of the LDA model denotes estimation of the probability that different sets of words (e.g., "experience," "sale," "price") appear in different pages of advertisements, conditional on the topic of the ad. Since each document page of advertisements contains a collection of words, the model will allow us to compute the probability that any one page of advertisements is comprised of job ads. Roughly put, the

Table 1: Results of the LDA Model

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
system	reg	car	day	street
experi	save	price	free	open
comput	store	new	new	inc
opportun	size	auto	one	ave
year	price	power	call	mass

Notes: The table provides the top five word stems for each of the five topics in our estimated LDA model performed on the *Boston Globe* display ad subsample.

model identifies sets of words that frequently appear together in the same documents within a text corpus. For example, if there were only two types of ads in our newspaper data, job ads or sales ads, one set of ads would be characterized by containing the words "experience," "years," or "opportunity." A second set of ads would be characterized by containing the words "store," "save," or "price."

To construct our LDA model, we take samples of pages of advertisements for each of our newspapers, separately: display ads in the *Boston Globe*, spanning 1960 to 1983; display ads in the *New York Times*, from 1940 to 2000; classified ads in the *Boston Globe*, from 1960 to 1983; classified ads in the *New York Times*, from 1940 to 2000; and classified ads in the *Wall Street Journal*, spanning 1940 to 1998.[8]

As an example of the partial output from our estimated model, Table 1 lists word stems which have the highest conditional probability of appearing in each of the five topics in our model. This model was estimated from pages of display ads appearing in the *Boston Globe*.[9] Words in Topic 1 tend to appear frequently in job ads, while other topics contain words prevalent in other types of advertisements.

At this point, each page is defined by a probability distribution over the set of possible topics. We pick the pages for which the likelihood of belonging to the topic that we a priori identify with vacancy postings is greater than 0.40. The choice of the cutoff balances a trade-off between throwing out too many vacancy postings (particularly job ads in sales-

[8]These time spans correspond to the original data provided to us by ProQuest. While the data set which we have posted online contains data starting in 1940, throughout this paper we restrict attention to job ads which were published beginning in 1960. We make this restriction since the focus of our paper is on the implications of occupations' task content on increasing labor income inequality, which began in earnest only in the 1960s.

[9]Computing the LDA model requires us to specify the number of topics present in the text. Since there is little a priori guidance in determining the optimal number of topics, for each source we choose the number of topics, one at a time, until we reach the largest number of topics which result in only one topic that is related to job vacancy postings. In Appendix C.1, we present the tables of predictive words for each newspaper in our sample.

related occupations) and including too many non-job ads in our data set.[10] Appendix C.1 provides additional details regarding the LDA procedure, in general, and the application of this procedure to our particular problem. See, also, Blei, Ng, and Jordan (2003) and Hoffman, Blei, and Bach (2010).[11]

Discerning the boundaries between vacancy postings and identifying each advertisement's job title. In the ProQuest data set, the advertisements on a single page are all grouped in a single text field. Our next task is to identify when one vacancy posting ends and a second posting begins. Here, certain phrases at the beginning or end of individual help wanted ads allow us to identify the boundaries between ads. In particular, we use a rule that relies on certain common patterns across ads and relies on the presence of i) street addresses, ii) certain phrases (e.g., "send [...] resume" or "equal opportunity employer"), and iii) the formatting of job titles. To search for job titles, we look for lines of text that are either i) all capitalized or ii) only one or two words appear. Within these lines, we search for words which appear in O*NET's "Sample of Reported Job Titles." If there is a match, we record the contents of the lines as the ad's job title.

Sample of results Table 2 presents the result of our procedure applied to the unprocessed text from Figure 1. At least in this small snippet of text, our algorithm is able to correctly parse job titles and identify the boundaries between the three advertisements. In identifying the boundary between the first and second ad, line breaks before "payments clerks" were helpful. In addition to these line breaks, an address—82 Devonshire Street, Boston MA 02109—helps identify the boundary between the second and third job posting. Finally, our spell-checker is able to fix some transcription errors. At the same time, even after processing, a few transcription errors remain. Moreover, because part of the text of the first job ad appears before the job title, it is not included with the remaining text for the "mutual funds clerks" position.

To sum up, our process allows us to transform unstructured text into a set of distinct job ads and undo a portion of the error induced by ProQuest's optical character recognition software. However, since our efforts to process the text are imperfect, our resulting data set will necessarily contain some measurement error. Still, as we will demonstrate in Section 3, given the massive number of vacancy postings which we have been able to process, our

[10] In practice, the model assigns low probabilities to intermediate likelihood values, and so the errors we make are relatively insensitive to our choice of cutoff.

[11] Within economics, we are aware of one application of LDA, which classifies Federal Open Market Committee statements by topic. See Fligstein, Brundage, and Schultz (2014) and Hansen, McMahon, and Prat (2014).

Table 2: Processed Text from the *Boston Globe*, November 4, 1979, Display Ad #133

job title	text
	rapid growth through continued innovation and diversification If you are highly motivated person who takes pride in your work and the company you work for consider career with Fidelity
mutual funds clerks	individuals with 1 2 years funds transfer experience to process Keogh IRA accounts and adjustments
payments clerks	varied Les processing new account applications and payments stir nt flt near fr NLRB ct t fl mnirk and maintaining client strong record keeping 50 wpm data control department sorting and brokerage environments with benefits package for our Boston convenient to the market gr Willie lu5lty success l 1 q1 a1 ol p 1sl1 vmplov fidelity group 82 devonshire street Boston ma 02109
terminal operators	position involves typing policy related information into computer terminal no previous computer experience required typing 5055 wpm excellent benefits plus work incentive program in addition to starting salary of 150 165

Notes: The table presents text from the first three of the 12 vacancy postings in a page of display ads in the *Boston Globe*.

database is still quite informative about occupations and their characteristics.

2.2 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 data, postings for the same occupation appear via multiple distinct job titles. For example, vacancy postings for registered nurses will be advertised using job titles which include "rn," "registered nurse," or "staff nurse." These job titles should all map to the same occupation — 291141 using the BLS Standard Occupational Classification (SOC) system.

From our list of job titles, we apply two complementary methods to group ads according to their SOC codes. First, for the most common 1000 job titles from our *Boston Globe*, *New York Times*, and *Wall Street Journal* job ads, we map the titles to their SOC codes using www.onetsocautocoder.com. These top 1000 job titles cover 2.2 of the 6.6 million ads from the previous subsection's processed newspaper data.

For the remaining ads, we apply a *continuous bag of words* (CBOW) model, in combination with an ancillary data set provided to us by EMSI, to attempt to identify the ad's SOC code.[12] Roughly put, this CBOW model allows us to find synonyms for words or phrases

[12]While onetsocautocoder.com is a useful, accurate tool to retrieve commonly occurring job titles, there are hundreds of thousands of unique titles occurring in at least two postings in our newspaper data, which

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 "nurse," "rn," and "staff nurse" all tend to appear next to words like "patient," "care," or "blood" one would conclude that "rn" and "nurse" have similar meanings to one another. For additional background on CBOW models, and details on our implementation, see Appendix C.3.

This CBOW model is useful since we have an ancillary data set containing a large correspondence between job titles and SOC codes. EMSI has provided us a data set of the text from online job ads, originally posted between April 2011 and March 2017. These ads contain a job title, a SOC code, and text describing the job characteristics and requirements. We use the 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. We use this model at different points in the paper to find synonyms for words or phrases. At this stage, our CBOW model allows us to identify the closest job title in the EMSI data set for each job title in our newspaper text. Since job titles in the EMSI data set have an associated SOC code, we can obtain the SOC code for any job title in our newspaper text. We retrieve these SOC codes on all of the job titles which appear at least twice in our newspaper text. In combination with the first approach, we have been able to collect SOC codes for 4.1 million job ads.

2.3 Eliciting Skill- and Task-Related Information

Within the body of our job ads, we will 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 not only job titles, but also these occupational characteristics into a manageable number of groups. We follow four 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 analytic*.[14] In our main application of these categories, we begin with the list of

makes manually retrieving them infeasible.

[13]Throughout this paper, we interpret the words as accurate representations of the positions the 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, where respondents may or may not accurately answer questions about their job's tasks (Autor, 2013). Our analysis is unaffected by level differences job descriptions' accuracy, and would only be affected by trends in the representation of jobs over time.

[14]The data set used in Spitz-Oener (2006) is a questionnaire given to West German workers. Building on

words related to each of her five tasks. For each task, we augment the list with words whose meanings are similar to those in the original list, where similarity is determined by a CBOW model. 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 will consider a narrower mapping between categories and words, one which only relies in Spitz-Oener (2006)'s definitions as enumerated in footnote 14.

We also consider three 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. First, with the aim of validating our data set, 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). For each O*NET Element, we begin by looking for words and phrases related to the O*NET Title and, refer to the O*NET Element Description to judge whether these synonyms should be included, as well as if other words should be included. As a second approach, we append to our initial lists of words and phrases an additional set of words, using our continuous bag of words model. Each of the two approaches, the "judgment-based" procedure and the "continuous bag of words model-based" procedure of identifying similar words, has its strengths and weaknesses. On the one hand, the first procedure is clearly ad hoc. Moreover, the continuous bag of words model has the advantage of accounting for the possibility that employers' word choice may differ within the sample period.[15] On the other hand, there is a danger that the bag of words model will identify words as synonymous even if they are not.

Second, for the purpose of mapping our new task measures to changes in wage inequality,

her mapping from survey question titles to task categories, we search for the following sets of words for each category: 1) nonroutine analytic: analyze, analyzing, design, designing, devising rule, evaluate, evaluating, interpreting rule, plan, planning, research, researching, sketch, sketching; 2) 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; 3) nonroutine manual: accommodate, accommodating, accommodation, renovate, renovating, repair, repairing, restore, restoring, service, serving; 4) routine cognitive: bookkeeping, calculate, calculating, correcting, corrections, measurement, measuring; 5) 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 the case: Use of the word "innovative" has increased more quickly than "creative" over the sample period. To the extent that our ad hoc classification included only one of these two words, we would be mis-characterizing trends in the O*NET skill of "Thinking Creatively." The advantage of the continuous bag of words model is that it will identify that "creative" and "innovative" mean the same thing because they appear in similar contexts within job ads. Hence, even if employers start using "innovative" as opposed to "creative" part way through our sample, we will be able to consistently measure trends in "Thinking Creatively" throughout the entire period.

we map our text to that in Firpo, Fortin, and Lemieux (2014). Firpo, Fortin, and Lemieux categorize work activities and contexts into five groups: Information Content, Automation/Routine, Face-to-Face Contact, On-Site Job, and Decision-Making.[16] Our last mapping is to skills in Deming and Kahn (2017)'s study of the relationship between firms' characteristics and the skill requirements in their vacancy postings, which we explore in the Appendix to provide complementary results.

2.4 Descriptive Statistics

Using the text from the *Boston Globe*, *New York Times*, and *Wall Street Journal*, our algorithm from Section 2.1 results in a data set with 6.6 million vacancy postings.[17] Among these vacancy postings, we have been able to retrieve a SOC code for 4.1 million ads. Table 3 lists the top occupations in our data set. The first two columns list common job titles among the 6.6 million total vacancy postings, while the last four columns present the top SOC codes. Across the universe of occupations, our newspaper data represents a broad swath of Management, Business, Computer, Engineering, Life and Physical Science, Healthcare, Sales, and Administrative Support occupations, but it under-represents Construction occupations and occupations related to the production and transportation of goods. See Appendices A and B for an analysis of the representativeness of our newspaper data relative to the decennial census and CPS. (These latter data sets we access via Ruggles, Genadek, Goeken, Grover, and Sobek, 2015.)

Table 4 presents, for each task in Spitz-Oener (2006)'s classification, the most task-intensive occupations. Occupations which are intensive in nonroutine analytic tasks are concentrated in Architectural and Engineering occupations (SOC codes beginning with "17") and Life, Physical, and Social Science Occupations (SOC codes beginning with "19"). Management ("11") and Sales ("41") occupations mention nonroutine interactive tasks frequently, while customer-service and maintenance related occupations (in a variety of SOC groups) have high nonroutine manual task intensities. Routine cognitive and routine manual task-related words are mentioned frequently in advertisements for clerical and production-related positions.

[16] See Appendix Table A.2 of Firpo, Fortin, and Lemieux (2014) for the mapping between their task groups and O*NET work activities and work contexts. Since our constructed data set includes only work activities, our corresponding measures of Firpo, Fortin, and Lemieux (2014)'s O*NET groups are based only off of work activity variables, and since the Automation/Routine task group is defined only by to O*NET work contexts, we will not be able to measure this variable.

[17] This 6.6 million figure excludes vacancy postings for which we cannot identify the job title or which contain a substantial portion, 35 percent or greater, of misspelled words. We also exclude ads with fewer than 15 words.

Table 3: Common Occupations

Job Title		6-Digit SOC Occupations		4-Digit SOC Occupations	
Description	Count	Description	Count	Description	Count
Secretary	117.3	436014: Secretary	275.6	4360: Secretary	408.6
Sales	55.4	439061: Office Clerk	182.2	4390: Oth. Admin. Support	340.3
Assistant	53.6	132011: Accountant	153.9	1320: Accountant	194.6
Accounting	51.0	412031: Retail Sales	149.8	1511: Computer Sci.	185.4
Clerk	50.0	433031: Bookkeeper	131.0	4330: Financial Clerks	174.8
Typist	50.0	439022: Typist	118.1	4120: Retail Sales	172.0
Salesperson	42.0	291141: Nurse	117.8	2911: Nurse	140.8
Engineer	41.6	111021: General Mgr.	91.6	1721: Engineers	132.1
Manager	41.3	172112: Ind. Engineer	88.1	1110: General Mgr.	127.7
Bookkeeper	40.5	113031: Financial Mgr.	68.1	1130: Financial Mgr.	101.8

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

3 Comparison of the New Data Set to O*NET

Before exploring trends in occupational task content and the implications for the distribution of earnings, we briefly compare our new data source to the widely used O*NET database. Our goal is to demonstrate the usefulness of newspaper text for measuring the task content of occupations.

Figure 2 plots the O*NET Importance measure and *Boston Globe*, *New York Times*, and *Wall Street Journal* keyword frequencies. Each panel plots this comparison for a separate O*NET Element. Within each panel, each data point represents a single 4-digit SOC occupation: For example, the "1920" (the code for Physical Scientists) in the top right panel indicates that the O*NET Importance of the skill of "Critical Thinking" is 4.0 (on a scale from 1 to 5), while in the newspaper data, we detect 0.11 Critical Thinking related keywords per ad. Figure 2 includes one work style (Independence), one skill (Critical Thinking), one knowledge requirement (Computers and Electronics), and one work activity (Interpreting for Others), each in a separate panel The correlations (weighted by the number of vacancy postings in our newspaper data) in these four plots are 29 percent, 59 percent, 54 percent, and 51 percent, respectively.[18] The relationship between our newspaper keyword frequencies and O*NET Importance measures appears to be nonlinear: For low-to-moderate O*NET Importance values, skill- and task-related keywords are very rarely mentioned in vacancy postings. Our algorithm detects skill- and task-related keywords only if the O*NET Impor-

[18]Across all 125 O*NET Elements, the unweighted correlations are lower by 7 percentage points on average.

Table 4: Top Occupations by Spitz-Oener (2006) Task Category

Nonroutine Analytic			Nonroutine Interactive		
1720: Engineers	69.3	0.89	1120: Sales Managers	70.1	1.11
1721: Engineers	132.1	0.89	4140: Sales Rep., Whole./Man.	64.1	0.80
1930: Social Scientists	5.2	0.70	4130: Sales Rep., Services	100.4	0.64
1910: Life Scientists	9.0	0.69	1110: Top Executives	127.7	0.61
1311: Business Operations	71.3	0.69	4120: Retail Sales	172.0	0.60
Nonroutine Manual			**Routine Cognitive**		
4930: Vehicle Mechanics	28.0	0.38	4330: Financial Clerks	174.8	0.22
4910: Maintenance Supervisors	14.1	0.29	4390: Other Admin. Support	340.3	0.10
4990: Other Maintenance	31.5	0.29	4320: Commun. Equip. Operators	11.0	0.09
4920: Electrical Mechanics	6.7	0.26	4310: Administrative Support	42.1	0.09
4340: Record Clerks	32.3	0.20	3730: Grounds Maintenance	1.4	0.08
Routine Manual					
5140: Metal and Plastic	32.0	0.14			
4990: Other Maintenance	31.5	0.07			
5141: Metal and Plastic	13.0	0.06			
4722: Construction Trades	2.6	0.05			
5120: Assemblers & Fabricators	9.8	0.05			

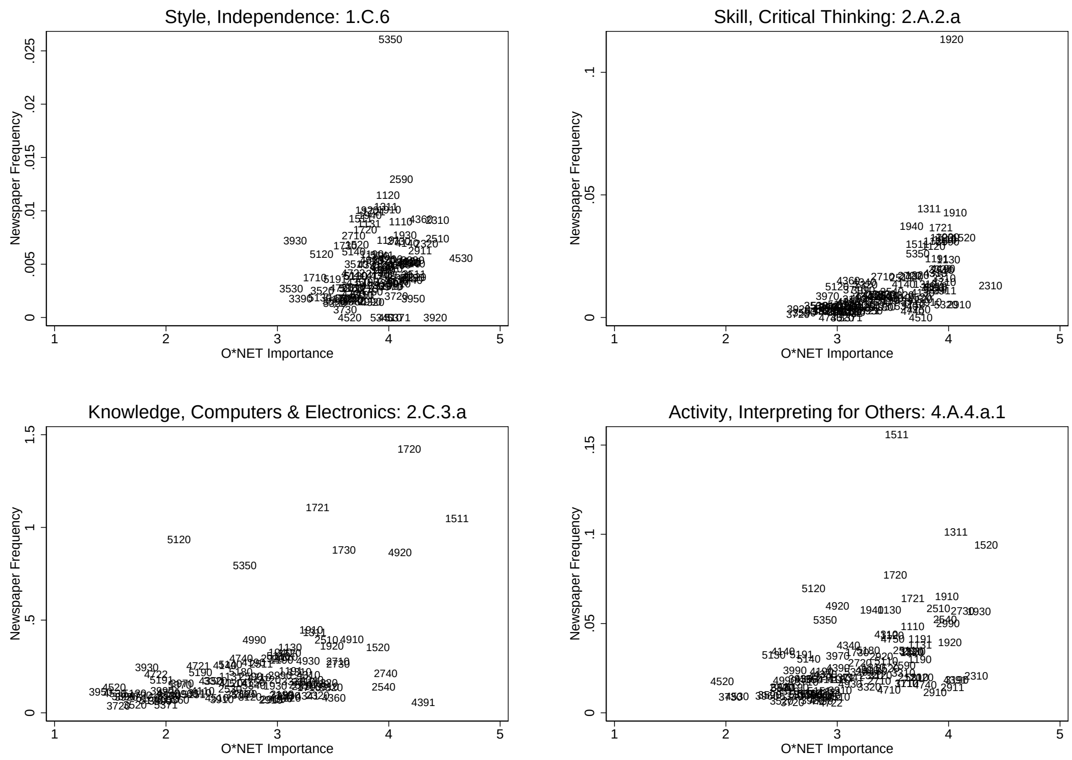
Notes: This table lists the top five 4-digit occupations according to the frequency with which different activity-related words are mentioned. Within each panel, the first column gives the SOC code and title; the second column gives the number of job ads in our data set (in thousands); and the final column gives the frequency (mentions per vacancy posting) of task-related words. Only occupations with at least 200 advertisements, representing 104 4-digit SOC codes, are included.

tance measures are sufficiently high, above 3 or 3.5.

The four relationships depicted in Figure 2 are broadly representative of the concordance between O*NET Importance measures and our vacancy postings' keyword frequencies: The correlation between our measures and existing O*NET measures of occupational work styles, skill, knowledge requirement, and activity measures are, for the most part, in the 0.40 to 0.70 range, and are somewhat higher for knowledge requirements, skills, and activities (where the mean correlations are 0.58, 0.54, and 0.50, respectively) than for work styles (where the mean correlation is 0.33).[19]

[19] While we view O*NET as a useful and reliable benchmark, it too has its issues, which extend beyond its inability to track within-occupation changes in job characteristics over long time horizons. In their review of the design of the O*NET data collection program, the National Research Council (2010) identify several aspects of O*NET which may limit its usefulness as a research tool. Summarizing these issues, Autor (2013) writes that both the Dictionary of Occupational Titles and O*NET "job content measures are often vague, repetitive, and constructed using ambiguous and value-laden scales that are likely to confuse respondents" (p. 191). We should neither expect nor hope that our measures exactly align with O*NET measures, but we interpret the correlations as reassuring evidence that the newspaper text is a valuable source of task data.

Figure 2: O*NET Importance Measures and Newspaper Keyword Frequencies



Notes: Each panel corresponds to one O*NET element. In each panel, each point represents a SOC code. The value of the x-axis represents the O*NET Importance measure (from version 20.1 of the O*NET database). The y-axis measures the number of keyword appearances per job ad, using data from the *Boston Globe*, *New York Times*, and *Wall Street Journal*.

4 Trends in Tasks

In this section, we document trends in occupational tasks from 1960 to 2000. The main results from these exercises are that i) between the 1960s and 1990s words related to non-routine interactive and analytic tasks have increased in frequency, while words related to routine manual and cognitive tasks have decreased in frequency, and ii) a large fraction of these changes are due to changes within occupations, rather than changes in employment shares across occupations. Lastly, we compare trends for three particular groups of occupations and show that our results resonate with previous findings in the literature.

4.1 Overall Trends

Table 5 present changes in task-related keywords, grouped according to the definitions introduced in Spitz-Oener (2006). The first column lists the task groups, while columns 2 through

Table 5: Trends in Keyword Frequencies

	Frequency								Within
	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-00	Share
Nonroutine	3.19	3.50	3.33	3.63	3.93	3.86	4.63	5.48	0.89
Analytic	[3.16]	[3.47]	[3.27]	[3.49]	[3.71]	[3.50]	[4.32]	[5.22]	(0.04)
Nonroutine	4.94	4.56	4.57	5.18	5.84	5.97	6.56	7.03	0.94
Interactive	[4.94]	[4.52]	[4.44]	[5.07]	[5.85]	[5.71]	[6.29]	[6.89]	(0.05)
Nonroutine	2.50	2.36	2.08	2.54	2.59	2.63	2.59	2.73	1.49
Manual	[2.52]	[2.40]	[2.04]	[2.56]	[2.65]	[2.67]	[2.54]	[2.86]	(0.68)
Routine	1.00	0.91	0.69	0.65	0.74	0.71	0.59	0.54	0.98
Cognitive	[0.99]	[0.90]	[0.67]	[0.61]	[0.73]	[0.76]	[0.61]	[0.54]	(0.04)
Routine	0.70	0.58	0.45	0.42	0.27	0.15	0.10	0.07	0.97
Manual	[0.71]	[0.60]	[0.51]	[0.50]	[0.31]	[0.18]	[0.11]	[0.10]	(0.04)

Notes: Occupations are defined using the 4-digit SOC classification. For each column, we compute average keyword frequencies (across occupations) for each five-year period. In these averages, occupations are weighted according to the number of workers in the decennial census, interpolating within-decade observations using counts from the 1960, 1970, 1980, 1990, and 2000 censuses. The numbers without square brackets represent the average keyword frequency in each five-year period. The numbers in square brackets represent the keyword frequencies which would prevail, holding fixed the number of workers in each occupation according to their values in 1960. The final "Within Share" column gives the ratio of the 1960-64 to 1995-2000 change according to the within-occupation component relative to the 1960-64 to 1995-2000 overall change. In this final column, the term within parentheses gives the bootstrapped (re-sampling ads from our newspaper text) standard errors of the "Within Share."

9 compute two measures of task content, broken down in five year periods. The measure at the top of each row, which we interpret as the aggregate task content in the economy, presents average mentions in each occupation using census employment shares as weights. According to this measure, mentions of nonroutine tasks increased between the early 1960s and late 1990s by 54 percent for analytic tasks (3.19 versus 5.48 mentions per thousand ad words), 35 percent for nonroutine analytic tasks, and 9 percent for nonroutine manual tasks. Conversely, the largest decline in keyword frequency occurred for words related to routine manual tasks, decreasing from 0.70 to 0.07 mentions per thousand ad words. The decline of routine cognitive tasks is also considerable, going from 1.00 to 0.54 mentions per thousand words.

These changes reflect both between-occupation and within-occupation changes in task-related keywords. 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:

$$\overline{T}_t - \overline{T}_{1960} = \sum_j \vartheta_{j,1960} \left(\widetilde{T}_{jt} - \widetilde{T}_{j,1960} \right) + \sum_j \left(\vartheta_{j,t} - \vartheta_{j,1960} \right) \widetilde{T}_{jt}. \tag{1}$$

In this equation, $\widetilde{T}_{jt}$ is a measure of task-related word frequency in postings for occupation j in year t. The ϑ_{jt} terms measure the share of workers in occupation j at time t, while $\overline{T}_t$ denotes the share-weighted average frequency of the task-related word at time t. On the right-hand side of Equation 1, the first sum represents the change in overall keyword frequencies due to within-occupation changes in task-related word mentions. The second sum represents the contribution of changes in the share of workers in different occupations on overall keyword frequencies. We use a 4-digit SOC classification to perform this decomposition separately for each of the five tasks introduced in Spitz-Oener (2006). The final column of Table 5 lists the proportion of the overall changes in keyword frequencies, comparing the 60-64 column to the 95-00 column, which are due to the within-occupation component. For every task, within-occupation changes account for a predominant share of the overall changes.

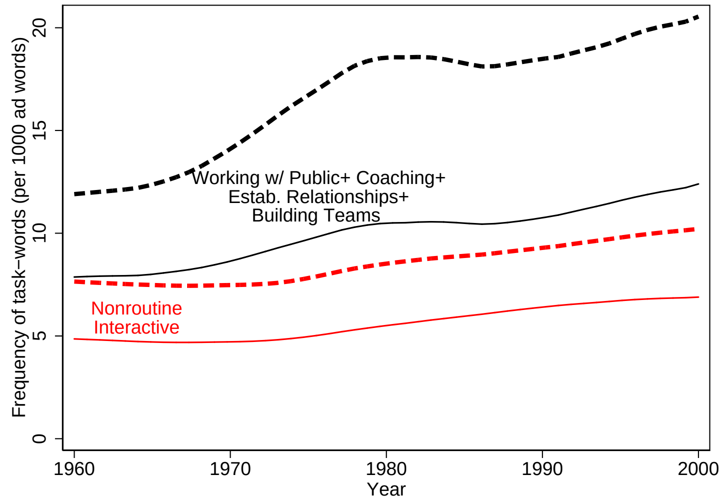
Sensitivity Analysis

In Appendix D we consider the sensitivity of the results given in Table 5 to the mapping of words to tasks, as well as to the level of detail of the occupation codes used.

First, we recompute Table 5 with Spitz-Oener (2006)'s original mapping between tasks and words (i.e., excluding the words which we have appended from our continuous bag of words model). Second, we consider two alternate measures of occupational characteristics, one introduced by Firpo, Fortin, and Lemieux (2014) and the other by Deming and Kahn (2017). Using each of these three alternate measures of occupational characteristics, the predominant share of the overall changes in occupational characteristics occurs within rather than between 4-digit SOC codes.

The extent to which between-occupation changes are responsible for overall changes in keyword frequencies may potentially be sensitive to the coarseness of occupation definitions. If occupations are coarsely defined, one will tend to estimate that between-occupation changes are relatively unimportant. In a second robustness check, we re-estimate Equation 1 first using 6-digit SOC codes to classify occupations and alternatively using individual job titles to classify occupations. Instead of a median within share of 0.97, which we obtain from averaging over the final columns of Table 5, one would arrive at a median within share of 0.99 with a 6-digit classification and 0.68 with a job-title-based classification. For all three classifications, within-occupation changes are the primary source of the overall shift in task content.

Figure 3: Task Measures: Managers and All Occupations



Notes: The figure above plots the average number of task-related word mentions for two different types of tasks, both for managerial occupations (dashed lines) and for all occupations (solid lines). We apply a local polynomial smoother. Managerial occupations are defined as those with a SOC code between 1100 and 1199.

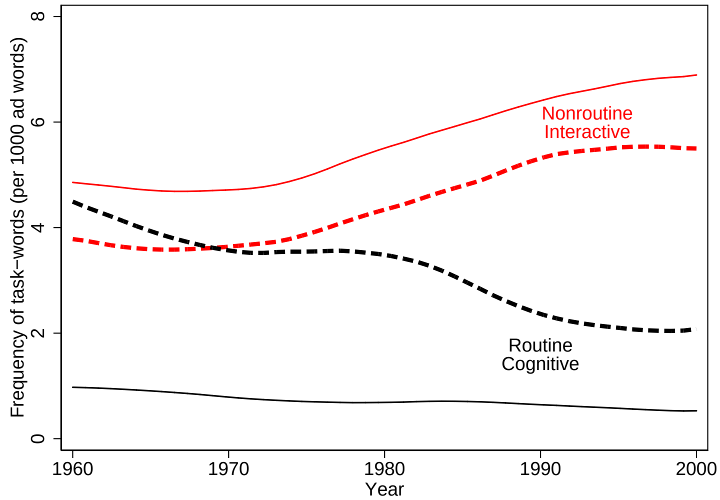
4.2 Narratives of Occupational Change

In this section, we present three vignettes that demonstrate how individual occupation groups have evolved over our sample period. Our goal is to provide concrete, illustrative examples of occupational change through the lens of our task and skill measures. We emphasize that these examples represent a portrait of long-run occupational change that was previously unobserved by researchers.

In Figure 3, we present two separate task measures for managerial occupations (in thick, dashed lines) and all occupations (in thin, solid lines). Between 1960 and 2000, the frequency of words related to nonroutine interactive tasks in managerial occupations increased by approximately one-third, from 7.5 to 10.0 mentions per 1000 ad words.[20] This measured trend aligns with the National Research Council (1999)'s characterization of the changing nature of managerial work. Summarizing the contemporaneous literature, the National Research Council (1999) write 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." (pp. 137-138) Motivated by this characterization, we also plot trends in the mentions of four O*NET work activities:

[20]While it is true that some of these changes reflect trends in the relative sizes of different managerial occupations, in Appendix D.3 we plot changes in task intensities within 4-digit SOC occupations. The results look quite similar.

Figure 4: Task Measures: Office Clerks and All Occupations

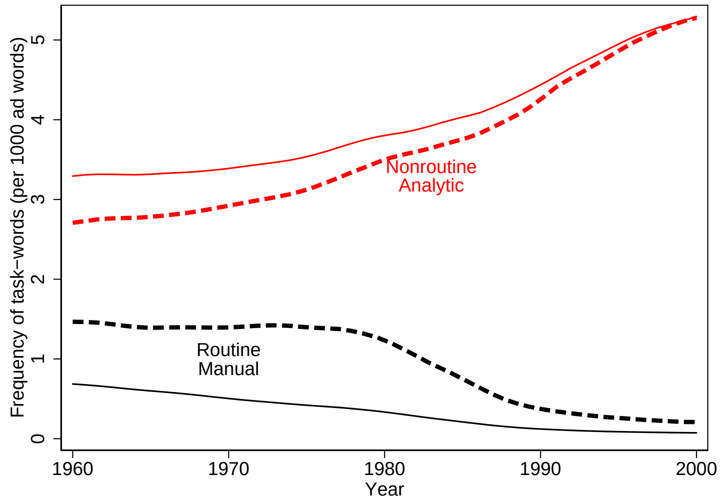


Notes: Office clerical occupations are those which have a SOC code between 4330 and 4341. We apply a local polynomial smoother. Averages across all occupations are depicted as the solid lines, and those across office clerical occupations are plotted as dashed lines.

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). Mentions of these four activities increased by more than 50 percent in managerial occupations, increasing more quickly than for the workforce as a whole. In sum, while interactive tasks have always been a requirement for managerial occupations, the importance of such tasks has widened since 1960.

Second, Figure 4 contrasts trends in task measures for office clerks compared to all occupations. While mentions of routine cognitive tasks has decreased for both office clerks and more generally for all occupations, the drop off has been more pronounced in clerical positions. Concurrently, the frequency of nonroutine interactive task-related words has increased in clerical ads roughly at the same pace as in other occupations. Both trends are consistent with Autor (2015)'s account of changes experienced by bank tellers (one particular office clerk occupation). He writes: "Increasingly, banks recognized the value of tellers enabled by information technology, not primarily as checkout clerks, but as salespersons, forging relationships with customers and introducing them to additional bank services like credit cards, loans, and investment products." (p. 7) (Unplotted, the frequency of nonroutine analytic-related keywords increased in office clerical occupations, partially closing the gap between these occupations and others.) Hence, much of the overall decline in routine cognitive tasks occurred in occupations like clerical occupations that were previously specialized in these tasks. The remaining jobs in these occupations increasingly contain nonroutine analytic and

Figure 5: Task Measures: Non-Supervisory Production Workers and All Occupations



Notes: Non-supervisory production occupations have an SOC code between 5120 and 5199. We apply a local polynomial smoother. Averages across all occupations are depicted as solid lines, and those across non-supervisory production occupations are plotted as dashed lines.

interactive tasks.

Finally, compared to the beginning of the sample period, the frequency of routine manual tasks has declined considerably, particularly in non-supervisory production occupations. Figure 5 presents these trends along with changes in the frequency of nonroutine analytic task related words. Nonroutine analytic task keywords have been increasingly mentioned in job ads for production workers. The trends in keyword frequencies in this figure are consistent with case studies of manufacturers' adoption of new information technologies (e.g., Bartel, Ichniowski, and Shaw, 2007). These new technologies substitute for workers who were previously performing routine manual tasks. The surviving production workers are those who have high levels of technical and problem-solving skills.

5 Implications for the Earnings Distribution

In this section, we use our new time-varying task measures to explore the implications of changes in occupations' task content for the earnings distribution. First, we bring our measures of occupations' task content to the decomposition methods of Fortin, Lemieux, and Firpo (2011). The goal of this exercise is to determine which occupational characteristics account for changes in the distribution of earnings between 1960 and 2000. Next, in Section 5.2, we develop a framework for interpreting occupations as a bundle of tasks. We embed this framework into a quantitative general equilibrium model of occupational sorting based on comparative advantage (akin to Heckman and Sedlacek, 1985, Heckman and Scheinkman, 1987, or more recently Burstein, Morales, and Vogel, 2015). We estimate this model and then use it to quantify the impact of changes in the demand for tasks on earnings inequality.

We view the two approaches as complementary methods for understanding the sources of increasing inequality, each approach with its own advantages and disadvantages. The statistical decompositions developed by Firpo, Fortin, and Lemieux (2014) are a flexible accounting device, helpful in determining whether observable characteristics of workers and occupations can account for changes in the earnings distribution. But it has some limitations. One is that the residual in the wage equation contains the contribution to wages of workers' sorting optimally across occupations. A second limitation resides in the interpretation of the decomposition in the face of general equilibrium adjustments. As noted by Firpo, Fortin, and Lemieux (2014), in response to a shock to the task demands, worker mobility across occupations will tend to limit changes in occupation prices, and this approach will attribute wage variation to changes in the returns to skills. Our Section 5.2 model is designed to grapple both with occupational sorting and general equilibrium effects. On the other hand, the model imposes parametric assumptions on workers' idiosyncratic ability to work in each potential occupation. Despite their differences, both approaches demonstrate that changes in occupations' task content generate a large fraction of the increase in 90-10 earnings inequality observed between 1960 and 2000.

5.1 Decompositions Using RIF Regressions

In this section, we perform a statistical decomposition due to Fortin, Lemieux, and Firpo (2011) to assess the role of occupations' task content on earnings inequality. Fortin, Lemieux, and Firpo introduce a method with which to decompose changes in the distribution of earnings across points in time on the basis of worker and occupational attributes. This method, which can be thought of as an extension of a Oaxaca-Blinder decomposition, breaks down changes over time in any quantile of the earnings distribution into the contribution of observable changes in worker and occupational characteristics (the "composition" effect) and the contribution of implicit changes in the rewards to those characteristics (the "wage structure" effect). One can further break down the composition and wage structure effects into the parts belonging to each of these characteristics (e.g. worker's educational attainment, occupational task content, etc.).

The basis of the decomposition is the following specification of workers' earnings as a function of their observed characteristics and the tasks of the occupations in which they work:

$$\log W_{ijt} = \log \pi_{0t} + \sum_{h=1}^{H} T_{hjt} \log \pi_{ht} + \sum_{k=1}^{K} \alpha_k \tilde{S}_{kg} + \log \epsilon_{ijt}. \tag{2}$$

In this equation, i is an individual, W_{ijt} is the individual's wage, and $\tilde{S}_{gk}$ is an observable skill characteristic k for an individual who is in group g. And, $\log \pi_{ht}$ is a regression coefficient that represents the (log) price of task h in period t.

To perform our decompositions, we compute recentered-influence-function (RIF) regressions that describe, for workers in each quantile, the relationship between real log earnings and different occupational and worker characteristics. In these RIF regressions, the coefficient estimates will vary by quantile; coefficient estimates for the 10th and 90th percentiles are given in Appendix E.1. The worker characteristics include a race indicator and an indicator for marital status, along with education and potential experience categorical variables.[21]

Our occupational characteristics are the measures based on Spitz-Oener (2006) and Firpo, Fortin, and Lemieux (2014) that we have discussed in Section 2.3.[22] Since the different task measures (e.g., routine cognitive versus nonroutine interactive) are scaled differently — not only because employers tend to systematically mention some types of words more frequently than others, but also because our algorithm may be better at detecting certain types of words than others — it is necessary to standardize our task measures. We normalize $T_{hjt} \equiv \left(\widetilde{T}_{hjt} - \mu_h\right)/\sigma_h$; here μ_h is the mean and σ_h is the standard deviation of keyword frequencies for task group h across years and occupations. Standardizing variables this way eliminates permanent cross task variation in mentions but retains the ability to do comparisons both across occupations at a point in time and within occupations across time.

[21] The education categories are Some High School; High School; Some College (which is the omitted group); College; and Post-graduate. The potential experience categories are defined by ten-year intervals (with individuals with 20 to 29 years of potential experience belonging to the omitted group.) These education and potential experience categories are more coarsely defined than in Firpo, Fortin, and Lemieux (2014). We adopt this classification to be consistent with our analysis in Section 5.2.

The sample includes only males with age between 16 and 65. We restrict the sample to male workers since the RIF regression decompositions are ill-equipped to handle the large increase in female labor force participation observed in the first half of our sample period. Appendix E.2 presents our decompositions for samples that include female workers. When females are included in the sample, the (unexplained) "wage structure" effects of our task measures are considerably larger in the 1960s, but similar in other decades. The composition effects are also similar to what we report here.

[22] An implicit assumption in our statistical decomposition is that changes in the level of task mentions represents a real change in the jobs' task content. An important consideration is whether firms' tendency to post information about task content (relative to other characteristics of the job or employer) has shifted over the last four decades of the twentieth century. We do not find any systematic trends in the overall share of job ad text mentioning worker activities, providing some reassurance that changes in the level of task mentions can be interpreted as changes in both the level of the task, and the share of that task, in the job's total task content.

Decompositions Using Spitz-Oener's (2006) Classification

We begin by grouping tasks into nonroutine analytic, nonroutine interactive, nonroutine manual, routine cognitive, and routine manual categories. We highlight the importance of our time-varying occupational measures by employing two sets of decompositions. In the first, T_{hjt} is fixed to its sample mean for each occupation-task combination (thus eliminating time variation), and in the second, T_{hjt} is allowed to freely vary across time within occupations.

Figure 6 presents our decompositions, together with the total change at each percentile of the distribution, for each decade between 1960 and 2000. The thick solid line shows total observed changes for each percentile. According to the top-left panel, average earnings increased between 1960 and 1970, with little change in earnings inequality. In the following decades, from 1970 to 2000, inequality increased more sharply, while average earnings increased less. To what extent can occupational characteristics account for these changes?

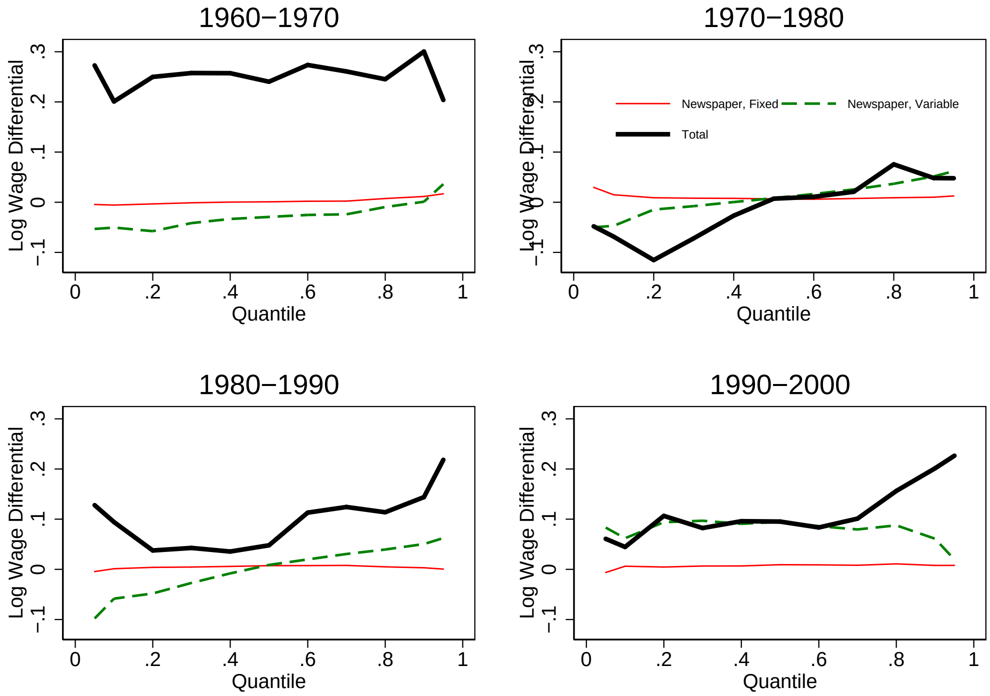
We plot the contribution of occupational task content, combining both the composition and wage structure effects, in the dashed and thin solid lines in Figure 6. The thin solid line — computed using task measures that are fixed throughout the sample — accounts for a small part of the changes in the earnings distribution. According to this measure, occupational characteristics account for a 2 percentage point increase in 90-10 inequality between 1960 and 1970 and no change in the 1970s, 1980s, or 1990s.

By contrast, when we allow task measures to vary across time, occupational characteristics account for a substantial increase in the median and 90-10 inequality. Between 1960 and 2000, with our preferred, time-varying measure of tasks, occupational changes can account for a large increase in male earnings 90-10 inequality: 5 percentage points in the 1960s, 10 percentage points in the 1970s, 11 percentage points between 1980 and 1990, and no change between 1990 and 2000.

In Figure 7, we sum up the decompositions in the four panels of Figure 6. Over these four decades, the 90-10 ratio of earnings inequality increased by 42 log points among male workers. Of this change in earnings inequality, our measures of occupational changes account for a 26 log point increase. In this figure, we also plot the two standard deviation confidence interval of the wage structure and composition effects. The two standard deviation confidence interval for the increase in 90-10 earnings inequality that is due to our task measures spans 4 to 48 log points.[23] On the other hand, the contribution of occupational characteristics to 90-10 earnings inequality is a modest 3 log point increase when using measures which

[23]These confidence intervals aim to measure the uncertainty surrounding our sampling of newspaper text. We construct 20 bootstrapped samples from our newspaper text. For each of these bootstrapped samples, we re-estimate our RIF regression based decompositions. To construct the dashed lines in Figure 7, we compute the standard deviation at each quantile of the RIF regression estimated contribution of our task measures to changes in the earnings distribution.

Figure 6: Decomposition of Real Log Earnings

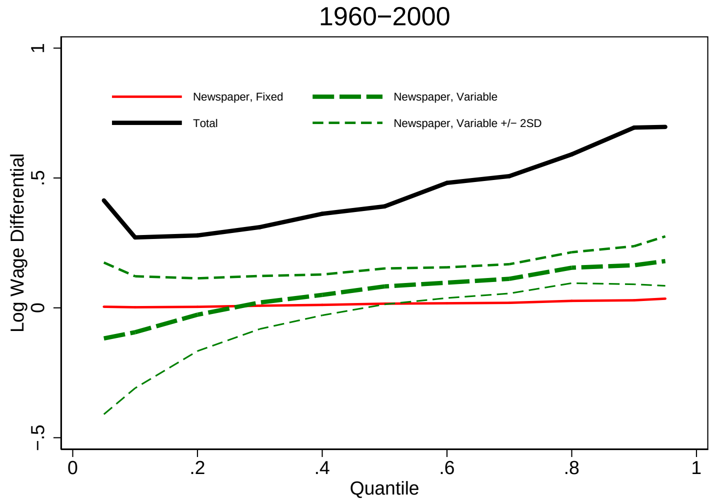


Notes: The thick solid line presents changes in log earnings of workers at different quantiles of the distribution. The thin solid line and the dashed line give the contribution of occupations (combining the composition effects and wage structure effects) using two different measures of occupational characteristics.

are fixed within occupations throughout the sample period. We conclude that variation in task content that includes changes within occupations across time can account for a large fraction of the rise in inequality. The contribution of occupational changes coming only from differences in task content between occupations is comparatively smaller.

To better understand the role of occupational changes in accounting for increased inequality, we plot the composition and wage structure effects separately for each newspaper-based measure (Figure 8). Occupational measures in which task measures are fixed yield modest composition and wage structure effects. Changes in the composition of tasks, as measured via fixed-over-time measures, account for a 6 percentage point increase in 90-10 inequality. On the other hand, the composition effect is substantially larger using our measure in which occupations' task measures vary as time progresses. In fact, for workers above the 90th percentile of the earnings distribution, these composition effects account for at least a 15 percentage point increase in earnings; for workers below the 10th percentile, the same com-

Figure 7: Decomposition of Real Log Earnings



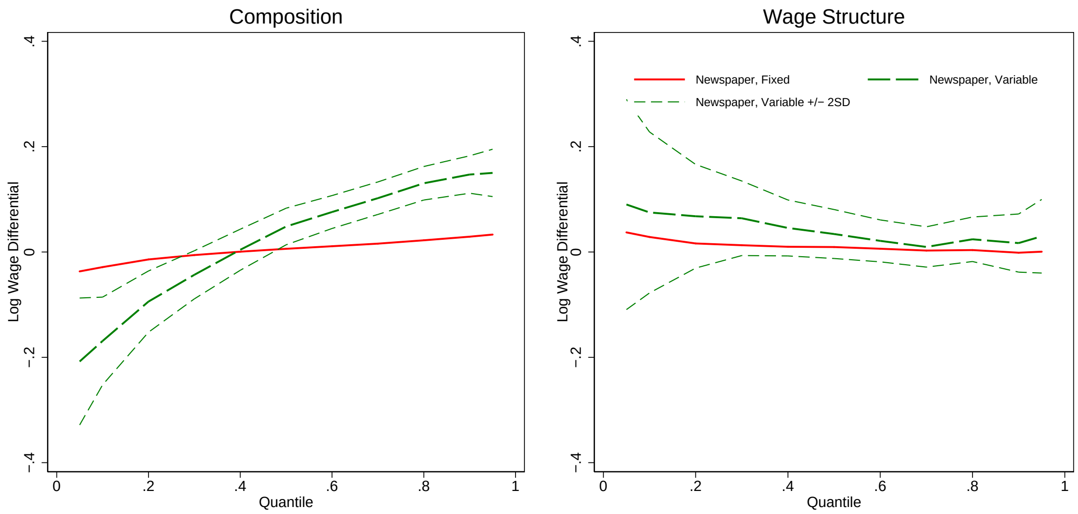
Notes: The thick solid line presents changes in log earnings of workers at different quantiles of the distribution. The thin solid line and dashed lines give the contribution of occupations (both through the composition effects and wage structure effects) using two different measures of occupational characteristics.

position effects account for at least a 17 percentage point decline. In sum, given the large changes in workers' activities within occupations that we uncovered in the previous sections, decompositions which rely only on between-occupation changes in task or skill-intensity understate the contribution of composition effects on changes in the earnings distribution.[24]

In Figure 9, we present the composition and wage structure effects, broken down by task. The key result from this exercise is that, among the estimated compositional effects, changes in routine manual tasks generate the largest changes in labor income inequality. Nonroutine interactive tasks also contribute to the increase in inequality, although to a lesser extent. These composition effects are due to a combination of i) a large decline in economy-wide routine manual and cognitive task content and a corresponding increase in nonroutine interactive and analytic task content; ii) a stronger association between workers' earnings and their occupations' tasks (a larger estimated $\log \pi$) for routine manual tasks at the bottom of the distribution than in the top; and iii) a stronger association between earnings and nonroutine interactive task content at the top of the distribution than in the bottom. Finally, changes in the "price" of tasks, as measured by the wage structure effects,

[24]There are analogous wage structure and composition effects for the demographic, experience, and education variables. We do not plot these here. Throughout the entire distribution, but especially across the right tail of the distribution, the composition effect for education is positive, indicating that increases in male workers' educational status is associated with higher earnings and earnings income inequality. These results are consistent with those of Firpo, Fortin, and Lemieux (2014).

Figure 8: Decomposition of Real Log Earnings: Composition and Wage Structure

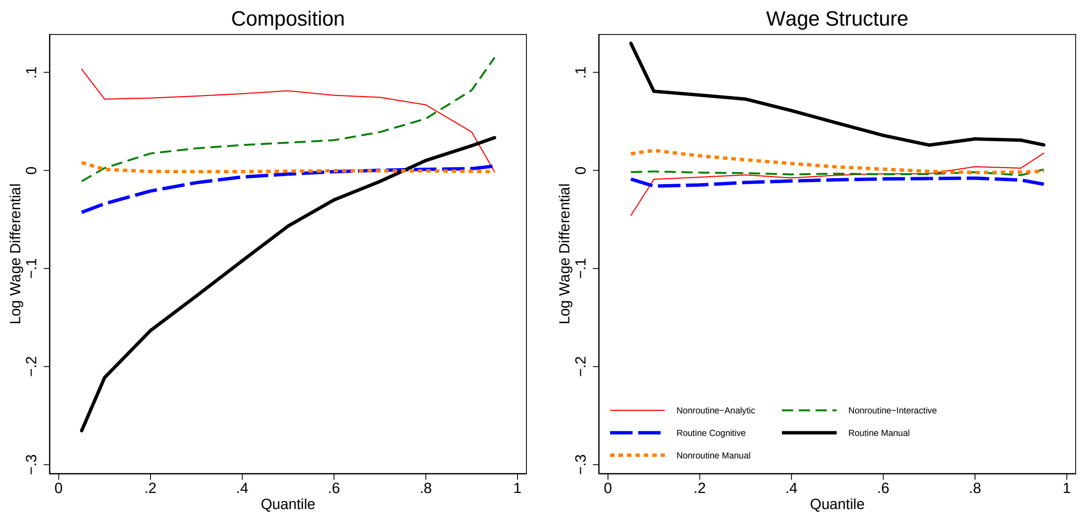


Notes: The left panel describes the contribution of occupational characteristics through compositional changes. The right panel describes the contribution of occupational characteristics through wage structure effects.

account for a 6 percentage point decline in 90-10 inequality.

The U.S. economy experienced several consequential changes since 1960, including dramatic technological progress and an increasing openness to trade. These developments have induced a shift in the composition of tasks which workers perform, and has re-shaped returns to these tasks. Our results suggest that changes in the task content of occupations translated shocks to trade and technology into widening inequality. Changes in the task returns played a more modest role. The implication is that, relative to the top of the distribution, the tasks which declined in importance over the sample period (routine manual tasks) were highly priced at the bottom of the wage distribution. Concurrently, tasks priced relatively highly for high-wage jobs (nonroutine analytic and interactive tasks) increased in importance between 1960 and 2000. A second implication follows from a comparison of the decompositions based on fixed and time-varying task content measures. With fixed task content, shifts in the tasks performed by workers can only arise via changes in the prevalence of occupations at different points of the distribution. Our decompositions indicate that such between-occupation shifts were of secondary importance in accounting for widening inequality, relative to the large within-occupation changes in task content that we have documented.

Figure 9: Detailed Decompositions of Real Log Earnings



Notes: The panels describe the contribution of individual occupational characteristics, through compositional changes (left panel) and wage structure effects (right panel) to changes in the earnings distribution. In this panel, we use the newspaper-based task measures which are allowed to vary within occupations across time.

Decompositions Using Firpo, Fortin, and Lemieux (2014)'s Classification

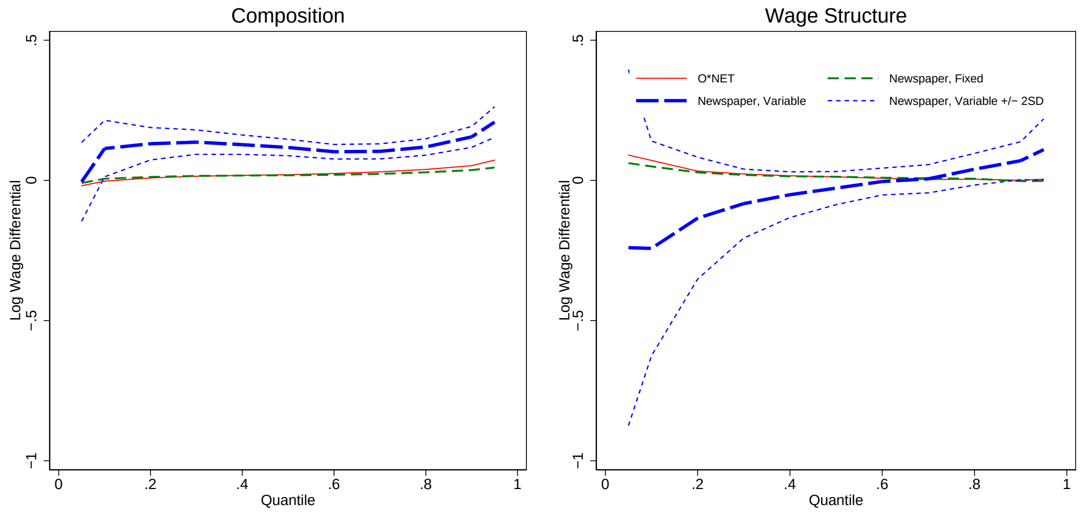
We now perform a similar set of decompositions using Firpo, Fortin, and Lemieux (2014)'s task classification. Figure 10, analogous to Figure 8, plots the composition and wage structure effects of our technology and offshoring measures on the earnings distribution between 1960 and 2000. Here, composition effects account for a 4 log point increase in 90-10 inequality. Wage structure effects account for a more substantial increase in inequality of 31 log points. In combination, Firpo, Fortin, and Lemieux (2014)'s task measures account for a 35 log point increase in 90-10 inequality.

Using fixed-over-time occupational measures yields much smaller results. Note that because there is a direct mapping between Firpo, Fortin, and Lemieux (2014)'s classifications and O*NET work activity measures, we can include measures based on O*NET. Results from decompositions based on our fixed-over-time newspaper task measures mirror those based on O*NET's measures, confirming that our newspaper-based measures track similar task content as those based on O*NET.

Sensitivity Analysis

In the appendices, we consider three sets of exercises to examine the sensitivity of our results. First, in Appendix E.2, we examine the robustness of our benchmark results to changes in the

Figure 10: Decomposition of Real Log Earnings: Composition and Wage Structure



Notes: The left panel describes the contribution of occupational characteristics through compositional changes using Firpo, Fortin, and Lemieux (2014)'s task definitions. The right panel describes the contribution of occupational characteristics through wage structure effects.

way we construct mappings — from newspaper text to Spitz-Oener's task groupings, to our transformation of the frequency of task-related words to T_{hjt}, and, finally, to the inclusion of women in the sample. Second, given the current discordance between the regions of the United States from which we are characterizing occupational characteristics (New York and Boston) and the regions we use to measure workers' income and demographic variables are measured (the entire U.S.), we re-compute Figures 7 through 10 using only census data from the New York and Boston MSAs (see Appendix E.3). Third, in Appendix E.4, we apply hourly wage data from the CPS Merged Outgoing Rotation Group, instead of annual labor income data from the decennial census, to analyze changes in income inequality.[25] These three different samples present a consistent depiction on the relationships between our job characteristics and trends in labor income inequality.

5.2 Decompositions in an Estimated Equilibrium Model

In this section we take an alternative approach to calculating changes in the earnings distribution. We begin with a quantitative model in which workers sort based on comparative advantage, along the lines of Burstein, Morales, and Vogel (2015). We consider the effect of

[25] We acknowledge a trade-off here: While we would prefer to use hourly wages as the basis for our measurement of income inequality, we wish to take advantage of the longer sample which the decennial census offers.

changes in the parameters of occupations' production functions, and interpret these changes as reflecting exogenous shifts in the demand for worker-produced tasks.

We consider a closed economy with many time periods $t = 1, 2, ...$ and a continuum of workers. Each individual worker i belongs to one of a finite number of groups, indexed by $g = 1, \ldots, G$. The mass of workers in group g is L_{gt}. Workers of type g are endowed with a set of skills S_{gh}, $h \in 1...H$, where h indexes different tasks which workers perform in their occupations. Worker i produces task h by combining skills and time: $q_{iht} = S_{gh} l_{iht}$. Here, l_{iht} equals the amount of time the worker allocates to task h. Each worker has a total of one unit of time, which she supplies inelastically in one of $j \in 1, ..., J$ occupations.

Worker i, if she works in occupation j, produces

$$V_{ijt} = \epsilon_{ijt} \prod_{h=1}^{H} \left(\frac{q_{iht}}{T_{hjt}} \right)^{T_{hjt}} \tag{3}$$

units of output. In Equation 3, ϵ_{ijt} is an occupation-worker unobserved efficiency level, and T_{hjt} gives the importance of task h in occupation j. We fix $\sum_h T_{hjt} = 1$ for each occupation and time period.[26] Variation in T_{hjt} across occupations captures, for example, that some occupations are more intensive in routine analytic tasks than others. To match our time-varying measures of task content, we allow the parameters T_{hjt} to evolve exogenously through time. Such changes could reflect, again as an example, declines in the demand for worker-performed routine manual or routine cognitive tasks caused by a decline in the price of computers, or increases in nonroutine interactive tasks related to the increased importance of team management.

We assume that worker i's wages, W_{ijt}, equal the value of the output she produces, $P_{jt} V_{ijt}$, which represents the product of the output produced and the price of the occupation j specific output. Taking P_{jt} as given, each worker chooses the amount of time spent on each of the H tasks to maximize the value of her output. The solution to this problem implies that $l_{ijt} = T_{hjt} \cdot \left(\sum_{h'=1}^{H} T_{h'jt} \right)^{-1} = T_{hjt}$.[27]

Plugging the optimal time allocation back into Equation 3 yields worker i's log wages

$$\log W_{ijt} = \log P_{jt} + \sum_{h=1}^{H} T_{hjt} \log S_{gh} + \log \epsilon_{ijt}. \tag{4}$$

[26]Throughout the model and its estimation, we adopt the stance that the five Spitz-Oener task categories proxy for the full set of mutually exclusive occupational tasks. A potentially omitted sixth task category would affect the estimation and results insofar as they represent a large share of worker time, and are orthogonal to — in their trends and skill requirements — those in the observed task groups.

[27]Due to the assumption that the production function of occupation-j specific output is Cobb-Douglas, individual abilities do not shape the time spent in each task, although they will affect occupational choice.

We assume that workers are freely able to choose their occupations to maximize their wages. Workers take as given each occupation-specific output price, and the task activities associated with each occupation. The idiosyncratic component of worker i's returns from working in occupation j, ϵ, is a Frechet-distributed random variable, drawn independently across occupations and workers, with shape parameter θ, i.e., $\Pr\left[\epsilon_{ijt} < z\right] = \exp\left(-z^{-\theta}\right)$. The parameter θ is an inverse measure of the dispersion of ϵ_{ijt} within group-occupation pairs.[28]

The Frechet formulation conveniently generates expressions for the fraction of group g workers which work in occupation j and the average wage of group g workers. The probability that a worker of type g sorts into occupation j is

$$\lambda_{gjt} = \frac{\left(P_{jt}\prod_{h=1}^{H}\left(S_{gh}\right)^{T_{hjt}}\right)^{\theta}}{\sum_{j'}\left(P_{j't}\prod_{h=1}^{H}\left(S_{gh}\right)^{T_{hj't}}\right)^{\theta}}. \tag{5}$$

Moreover, the average wages of group g workers equals

$$\overline{W}_{gt} = \Gamma\left(1-\frac{1}{\theta}\right)\cdot\left(\sum_{j}\left(P_{jt}\prod_{h=1}^{H}\left(S_{gh}\right)^{T_{hjt}}\right)^{\theta}\right)^{1/\theta}, \tag{6}$$

where $\Gamma\left(\cdot\right)$ is the Gamma function.

To close the model, we assume that there is a representative consumer who has CES preferences over the output produced by the different occupations:

$$V_t = \left[\sum_{j=1}^{J}\left(\sigma_j\right)^{\frac{1}{\rho}}\cdot\left(V_{jt}\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho-1}{\rho}},$$

with V_{jt} representing the sum of the output of individual workers, i, who produce occupation j output. Here, σ_j parameterizes the consumer's preferences for occupation j output, and ρ gives the preference elasticity of substitution across occupations' output. Given these preferences, we can write the market-clearing condition for occupation j output, equating expenditures on the occupation's output to the wage bill of workers who are employed in

[28] In Appendix F, we outline the extra assumptions necessary to derive from Equation 2, the basis of our statistical decompositions, from Equation 4.

occupation j:

$$\underbrace{\frac{\sigma_j \cdot (P_{jt})^{1-\rho}}{\sum_{j'} \sigma_{j'} (P_{j't})^{1-\rho}}}_{\text{(A)}} \underbrace{\sum_g \overline{W}_{gt} L_{gt}}_{\text{(B)}} = \sum_{g=1}^{G} \underbrace{\Gamma\left(1-\frac{1}{\theta}\right)\left(\sum_{j'}\left(P_{j't}\prod_{h=1}^{H}(S_{gh})^{T_{hj't}}\right)^{\theta}\right)^{1/\theta}}_{\text{(C)}} \underbrace{\lambda_{gjt} L_{gt}}_{\text{(D)}}. \tag{7}$$

The left-hand side of Equation 7 contains two terms: the share of total expenditures on occupation j output (term A), multiplied by total expenditures (term B). On the right-hand side, the total wage bill of workers in occupation j is computed as the sum of the wage bill of group g workers in occupation j; which in turn equals the product of the average wage of group g workers (term C) and the number of group workers (term D).

An equilibrium of our economy is the solution to $\{\lambda_{ij}\}$ and $\{P_j\}$, as determined by Equations 5 and 7.

With this model, we evaluate the impact of change in the demand for tasks on the equilibrium wage distribution. To do so, we first use data from a single period (1960) in combination with Equations 5 and 6 to estimate demographic groups' skills in producing tasks. Then, with our measures of changes in occupations' task content, we compute the counterfactual change in the distribution of wages, between 1960 and 2000, that is due only to changes in the demand for tasks. Appendix F details the equilibrium of the model, and delineates our algorithm to recompute the equilibrium to in response to changes in the Ts. Throughout the remainder of the section, we fix $\theta = \rho = 1.78$, following Burstein, Morales, and Vogel (2015).

In estimating our model, we need to map task-related keyword frequencies, as observed in the newspaper data ($\tilde{T}_{hjt}$), to our production function parameters (T_{hjt}). To ensure that our production function parameters sum to one for each occupation-year, we apply the following transformation: $T_{hjt} = \tilde{T}_{hjt} / \left(\sum_{h'} \tilde{T}_{h'jt}\right)^{-1}$.

Estimation of task-related skills

For each task h, we parameterize the relationship between skills and educational and demographic observables as follows:

$$\log S_{gh} = a_{h,gender} \cdot D_{gender,g} + a_{h,edu} \cdot D_{edu,g} + a_{h,exp} \cdot D_{exp,g}. \tag{8}$$

Here, $D_{gender,g}$, $D_{edu,g}$, and $D_{exp,g}$ are dummies for gender, education and experience, which define the categories g. We estimate these a parameters via a method of moments procedure.[29]

Our estimation recovers $a_{h,gender}$, $a_{h,edu}$, and $a_{h,exp}$ to minimize

$$\sum_g \sum_j \omega_{gj}^{\lambda} \left(\log \lambda_{gj,1960} - \log \lambda_{gj,1960}^{\text{data}}\right)^2 + \sum_g \omega_g^W \left(\log \overline{W}_{g,1960} - \log \overline{W}_{g1960}^{\text{data}}\right)^2 . \tag{9}$$

In Equation 9, $\lambda_{gj,1960}^{\text{data}}$ and $\log \overline{W}_{g1960}^{\text{data}}$ refer to the observed occupational shares and average wages.[30] The ω_{gj}^{λ} and ω_g^W are weights, characterizing the relative importance of each moment in our estimation. We compute these weights as the inverse of the variance of the moment (in the decennial census) across 40 bootstrapped samples.

Table 6 quantifies the patterns of comparative advantage across groups and tasks. For example, relative to other groups, workers with higher educational attainment tend to have a comparative advantage at performing nonroutine analytic tasks; high school graduates and workers with some college tend to have a comparative advantage in routine cognitive tasks; and high school dropouts have an advantage in routine and nonroutine manual tasks. Also, more experienced workers are more skilled in performing nonroutine interactive, and less skilled in routine cognitive tasks.

With the aim of substantiating these estimates, we make two points. First, as an assessment in-sample fit of the model, we compare $\log \lambda_{gj,1960}$ and $\log \lambda_{gj,1960}^{\text{data}}$ (across the 3240 group-occupation observations), and $\log \overline{W}_{g,1960}$ and $\log \overline{W}_{g1960}^{\text{data}}$ (across the 40 groups). The correlation between the model-estimated occupational shares and the observed shares is 0.52, while the correlation between $\log \overline{W}_{g,1960}$ and $\log \overline{W}_{g1960}^{\text{data}}$, is 0.95. Second, identification of the parameters in Table 6 follows transparently from workers' sorting patterns. Based on our set-up, the share of workers will be high for occupations in which their skills will be used intensively. For example, mirroring the estimates in Table 6, the correlation between the average number of words related to routine cognitive tasks in occupation j and an occupation-group's $\log \lambda_{gj,1960}$ is substantially higher for women compared to men (0.34 versus 0.00); for high school graduates compared to post-graduates (0.17 versus 0.10); and for workers with fewer than 10 years of experience compared to workers with 30+ years of experience (0.14 versus 0.09).[31]

[29] As in the previous subsection, we have two gender, five educational, and four experience groups, with "male" as the omitted gender category, "Some College" as the omitted educational category, and 20-29 years as the omitted experience category.

[30] Our estimation relies on data from 3240 ($=40 \cdot 81$) moments—representing information from 40 groups (2 genders, 5 educational categories, and 4 experience categories) and 81 occupations—to identify 121 parameters: 81 occupation-fixed effects plus 40 a coefficients.

[31] As a second example, the correlation between the frequency of words related to nonroutine analytic tasks and $\log \lambda_{gj,1960}$ is increasing across the five education groups: -0.31, -0.18, -0.01, 0.30, 0.39.

Table 6: Estimates of Skills

	Nonroutine Analytic	Nonroutine Interactive	Nonroutine Manual	Routine Cognitive	Routine Manual
Gender					
Female	-1.384	-0.473	0.269	3.472	-7.420
	(0.004)	(0.004)	(0.005)	(0.006)	(0.017)
Education					
< HS	-1.981	-1.155	2.628	-2.199	4.461
	(0.007)	(0.005)	(0.011)	(0.012)	(0.034)
High School	-0.785	-0.498	0.905	-0.006	2.757
	(0.008)	(0.005)	(0.011)	(0.010)	(0.033)
College	1.449	0.106	1.414	-1.002	-8.438
	(0.008)	(0.006)	(0.017)	(0.016)	(0.057)
Post-Graduate	1.394	-0.215	3.222	-2.778	-7.300
	(0.008)	(0.007)	(0.011)	(0.020)	(0.059)
Experience					
0-9 Years	-0.068	-0.602	-0.365	0.586	-2.560
	(0.005)	(0.005)	(0.007)	(0.010)	(0.020)
10-19 Years	-0.002	-0.219	0.216	0.354	-0.906
	(0.006)	(0.005)	(0.007)	(0.012)	(0.018)
30+ Years	-0.063	0.153	-0.026	0.239	-0.767
	(0.006)	(0.004)	(0.007)	(0.011)	(0.020)

Notes: To compute the standard errors, we re-sampled 40 times from the 1960 decennial census. For each bootstrapped sample, we recomputed the empirical occupational shares and group wages and then found the combination of parameters that minimized Equation 9. The omitted demographic groups are males, workers with Some College, and workers with 20-29 years of potential experience.

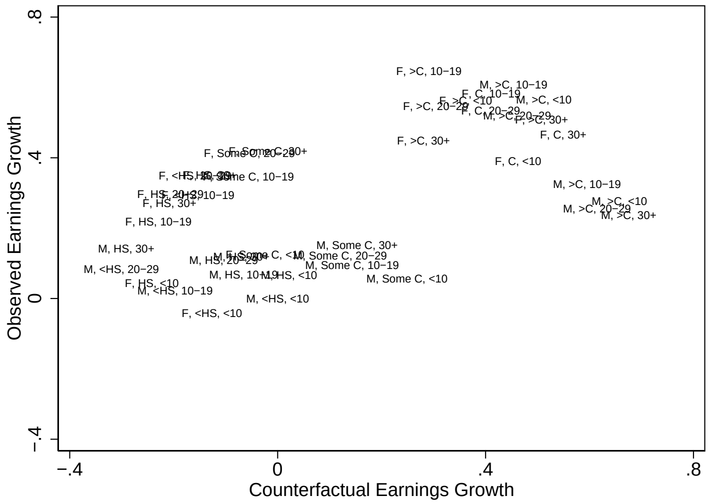
Counterfactual earnings distribution

Having estimated the distribution of skills, we now compute the counterfactual dispersion in wages that would otherwise have been obtained in succeeding decades—in 1970, 1980, 1990, and 2000—as a result of changes in the demand for tasks over our sample period. To do so, we consider an exogenous change in the T_{hjt} production function parameters (which, again, we measure through the changes in task-related keyword frequencies). Stemming from these changes in the demand for tasks, we compute the counterfactual equilibrium that would have obtained, fixing $\log S_{gh}$ to the values we estimated in the previous subsection, but allowing T_{hjt} to change over time.[32]

Figure 11 plots the growth in $\overline{W}_{gt}$ between 1960 and 2000, both as observed in the

[32] In the exercises in this section, we also fix the sizes of the 40 demographic groups to their 1960 values. An alternate procedure would allow these labor supplies to vary throughout the sample according to observed demographic changes. This alternate procedure yields similarly large increases in 90-10 inequality in the counterfactual equilibrium, compared to those reported in Figure 12.

Figure 11: Counterfactual and Observed Earnings Growth, 1960-2000, By Demographic Group



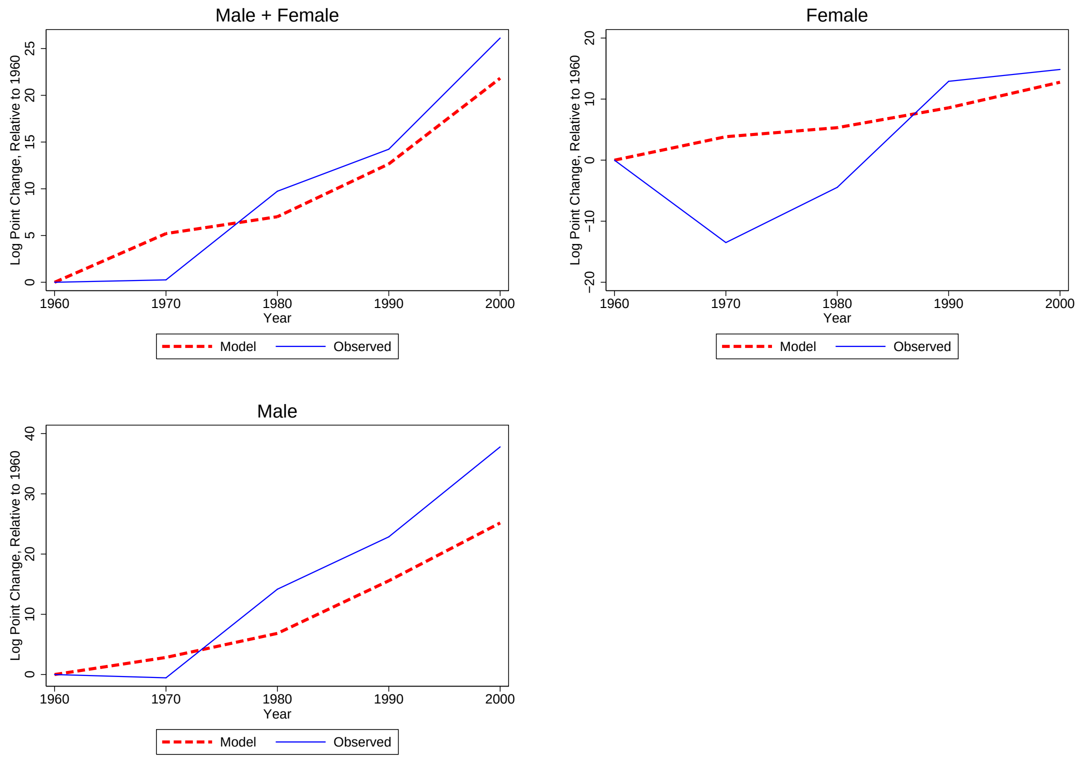
Notes: Each point gives the growth in wages for one of the 40 g groups The first character—"M" or "F"—describes the gender; the second set of characters—"<HS," "HS," "Some C," "C," or ">C"—the educational attainment; and the third set of characters the number of years of potential experience for the demographic group.

data and according to our counterfactual exercise. Wages increase most quickly for workers with college degrees or with post-graduate education. According to our estimated model, wage growth is fastest for high-education demographic groups because these groups have an advantage in nonroutine analytic and interactive tasks as of 1960, and the relative demand for these tasks has increased over the subsequent 40 years.

Increases in between-group inequality signify larger overall inequality. Figure 12 plots changes in 90-10 inequality, both the observed growth rate and the counterfactual growth rates that are due only to changes in T_{hjt}. Between 1960 and 2000, 90-10 inequality increased by 26 log points for the entire population, 38 log points for male workers, and 15 log points for female workers. According to our counterfactual exercises, within occupation changes in task demand account for much of the observed increase in 90-10 inequality: 22 log points for the entire sample, 13 log points for female workers, and 25 log points for male workers.

Discussion Occupational choice and wage data from the beginning of our sample suggest that workers with certain demographic characteristics (e.g., workers with a college degree) have an advantage in the tasks which have happened to grow over the subsequent forty years. Since these demographic groups were already highly remunerated in 1960, changes in the demand for tasks have increased earnings inequality.

Figure 12: Counterfactual and Observed 90-10 Inequality, 1960-2000



Notes: Each panel plots the change in 90-10 earnings inequality, as observed in the decennial census (solid line), and are due only to changes in the demand for tasks (dashed line).

Where do these shifts in the demand for tasks come from? Our model omits capital as an input in the production of tasks. Likely, some of the differential growth rates of T_{hjt} across task groups epitomize, at a more fundamental level, changes in capital prices. To the extent that the elasticity of substitution between workers and capital is relatively high in the production of routine tasks, a decline in the price of capital will be equivalent to a reduction in the demand for worker-performed routine tasks. Moreover, insofar as employers' innovation and technology adoption decisions are a function of the supply of workers of different skill types, changes in the T_{hjt} not only embody exogenous changes in the demand for tasks (which has been our interpretation here), but also reflect changes in the skill composition of labor supply. An interesting topic for future work would be to build on our model of occupations as a bundle of tasks, endogenizing changes in the T_{hjt} while acknowledging that workers and capital combine to perform each individual task and that innovations which lead to these declines in capital prices are potentially a response to the changing composition of the workforce.

The statistical decomposition in the previous section and the counterfactual exercise in this section, while markedly different in their approach and assumptions, result in comparable assessments of the role that task measures play in explaining increasing inequality. On the one hand, in contrast to the statistical decomposition, worker sorting plays a key role in the equilibrium model. Rewards to skills are occupation specific (and determined by the occupations' associated task content). Sorting amplifies the effect of observed changes in tasks on inequality.[33] Moreover, while heavily parameterized, the model takes into account the labor responses to changes in task prices induced by the shifts to the T_{hjt}. These general equilibrium effects also tend to generate more inequality explained by the model relative to the statistical decomposition.

On the other hand, in the model, we only evaluate the effect of within-occupation changes in task demands, while other shocks to fundamentals could also shape the income distribution. These other shocks include changes in the demand for occupations (measured by σ_j in the model), and potentially represent a source of widening income inequality that is encapsulated by the statistical decompositions and not the counterfactual exercises. Moreover, to the extent that the price of individual tasks vary within occupations, the RIF regressions will capture a source of inequality that the model cannot.

6 Conclusion

In this paper, we chronicle the changes in the U.S. occupational structure between 1960 and 2000. We document that a predominant share of changes in the skill and task composition of the workforce has occurred within rather than between occupations. Not only is it the case that occupations intensive in routine tasks have declined as a share of the workforce — a central pattern of the existing task literature — but also individual occupations' routine task content has declined as well.

We then demonstrate that within-occupation task changes are fundamentally important in an examination of the evolution of income inequality. We first perform a statistical decomposition on the earnings distribution and find that compositional changes associated with occupational characteristics can explain a 26 log point increase in labor income inequality (as measured by the 90-10 ratio) between 1960 and 2000. We then construct and estimate a Roy model in which changes in the demand for tasks determine workers' comparative advantage. Our estimated model accords with the results from our statistical accounting exercise.

[33] In principle, sorting on unobservable characteristics could lead the statistical decomposition to either overstate or understate the role of tasks in accounting for increased inequality. If, over time, there is increased sorting of workers with high unobserved ability into jobs at the top of the distribution, then the decomposition's wage structure effects will under-represent the contribution of task price changes to inequality. Firpo, Fortin, and Lemieux (2014) argue that this is likely the case. See their Section I.B.

Beyond this project, our newspaper data have the potential to address other economic questions related to the labor market. For example, in ongoing work we use the text to measure the adoption of new computer technologies, to study how these technologies interact with the task content of jobs. More generally, we view our newspaper-based job vacancy text as offering an opportunity to study questions that have been examined using online job vacancy text over a longer time horizon than previously possible (Hershbein and Kahn, 2016; Modestino, Shoag, and Ballance, 2016).

References

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

Autor, D. H. (2013): "The "Task Approach" to Labor Markets: An Overview," *Journal of Labor Market Research*, 46(3), 185 – 199.

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

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

Autor, D. H., L. Katz, and M. Kearney (2005): "The Polarization of the U.S. Labor Market," *American Economic Review*, 96(2), 189–194.

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

Bartel, A., C. Ichniowski, and K. 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), 1721–1758.

Bengio, Y., R. Ducharme, P. Vincent, and C. Jauvin (2003): "A Neural Probabilistic Language Model," *Journal of Machine Learning Research*, 3, 1137–1155.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003): "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Burstein, A., E. Morales, and J. Vogel (2015): "Accounting for Changes in Between-Group Inequality," Working Paper 20855, National Bureau of Economic Research.

Deming, D., and L. B. Kahn (2017): "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals," *Journal of Labor Economics*, (forthcoming).

Deming, D. J. (2017): “The Growing Importance of Social Skills in the Labor Market,” *Quarterly Journal of Economics*, (forthcoming).

DeVaro, J., and O. Gurtler (2017): “Advertising and Labor Market Matching: A Tour Through the Times,” *Journal of Labor Economics*, (forthcoming).

Firpo, S., N. M. Fortin, and T. Lemieux (2014): “Occupational Tasks and Changes in the Wage Structure,” IZA Discussion Papers 5542, Institute for the Study of Labor (IZA).

Fligstein, N., J. S. Brundage, and M. Schultz (2014): “Why the Federal Reserve Failed to See the Financial Crisis of 2008: The Role of Macroeconomics as a Sense making and Cultural Frame,” Institute for research on labor and employment, working paper series, Institute of Industrial Relations, UC Berkeley.

Fortin, N., T. Lemieux, and S. Firpo (2011): “Decomposition Methods in Economics,” *Handbook of Labor Economics*, 4, 1–102.

Gentzkow, M., and J. M. Shapiro (2010): “What Drives Media Slant? Evidence From U.S. Daily Newspapers,” *Econometrica*, 78(1), 35–71.

Hansen, S., M. McMahon, and A. Prat (2014): “Transparency and Deliberation within the FOMC: a Computational Linguistics Approach,” CEPR Discussion Papers 9994, C.E.P.R. Discussion Papers.

Heckman, J. J., and J. Scheinkman (1987): “The Importance of Bundling in a Gorman-Lancaster Model of Earnings,” *Review of Economic Studies*, 54(2), 243–255.

Heckman, J. J., and G. Sedlacek (1985): “Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market,” *Journal of Political Economy*, 93(6), 1077–1125.

Hershbein, B. J., and L. B. Kahn (2016): “Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings,” Upjohn Working Papers and Journal Articles 16-254, W.E. Upjohn Institute for Employment Research.

Hoberg, G., and G. Phillips (2016): “Text-Based Network Industries and Endogenous Product Differentiation,” *Journal of Political Economy*, 124(5), 1423–1465.

Hoffman, M. D., D. M. Blei, and F. R. Bach (2010): “Online Learning for Latent Dirichlet Allocation,” in *Advances in Neural Information Processing Systems 23: 24th Annual Conference on Neural Information Processing Systems 2010. Proceedings of a meeting held 6-9 December 2010, Vancouver, British Columbia, Canada.*, pp. 856–864.

Marinescu, I., and R. Wolthoff (2016): “Opening the Black Box of the Matching Function: the Power of Words,” NBER Working Papers 22508, National Bureau of Economic Research, Inc.

Michaels, G., F. Rauch, and S. J. Redding (2016): “Tasks and Technology in the United States 1880-2000,” Discussion paper, National Bureau of Economic Research.

Mikolov, T., K. Chen, G. Corrado, and J. Dean (2013a): "Efficient Estimation of Word Representations in Vector Space," Unpublished working paper.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013b): "Distributed Representations of Words and Phrases and Their Compositionality," in *Advances in neural information processing systems*, pp. 3111–3119.

Modestino, A. S., D. Shoag, and J. Ballance (2016): "Upskilling: Do Employers Demand Greater Skill When Skilled Workers are Plentiful?," Unpublished working paper.

National Research Council (1999): *The Changing Nature of Work: Implications for Occupational Analysis*. National Academy Press.

——— (2010): *A Database for a Changing Economy: Review of the Occupational Information Network (O*NET)*. National Academy Press.

Rehurek, R., and P. Sojka (2010): "Software Framework for Topic Modelling with Large Corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pp. 45–50, Valletta, Malta. ELRA.

Ross, M. (2017): "Routine-Biased Technical Change: Panel Evidence of Task Orientation and Wage Effects," *Labour Economics*, 48, 198–214.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015): "Integrated Public Use Microdata Series: Version 6.0," Minneapolis, MN: Historical Census Projects, University of Minnesota.

Spitz-Oener, A. (2006): "Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure," *Journal of Labor Economics*, 24(2), 235–270.

Yamaguchi, S. (2012): "Tasks and Heterogeneous Human Capital," *Journal of Labor Economics*, 30(1), 1–53.

A Additional Comparisons of the Newspaper Data to Existing Data Sources

Certain variables are present in both our new newspaper data and in previously available census and CPS data. With the aim of demonstrating the overall reliability of our newspaper data, we compare the frequency of different occupations, as well as educational characteristics for each occupation, across the two data sources. Unlike the comparisons that we made in Section 3, the occupational measures potentially vary across time in the data set to which we are comparing our newspaper data.

First, Figures 13 and 14 depict the share of workers (in the decennial census) in different occupational groups, along with the frequency of job ads in the same groups. Figure 13 presents this relationship at the 2-digit level. In 1960, 1980, and 2000, the correlations between census job frequencies and newspaper job ad data are 0.59, 0.78, and 0.65, respectively. Our newspaper data set over-represents the Sales, Health Practitioner, and Architecture/Engineering occupational groups, and conversely under-represents the Transportation, Production, and Installation and Maintenance occupational groups. (Hershbein and Kahn, 2016's data set of online job postings also exhibited a similar under-representation of blue-collar occupations.) Figure 14 presents the same set of relationships, now using a 4-digit SOC classification. Here, the correlations among the two measures of occupational size are somewhat weaker, ranging between 0.35 in 1960 to 0.49 in 1980.

Second, we compare measured educational attainment of occupations' workers in the decennial census to our vacancy postings' stated education requirements. In the newspaper text, we search among a list of acronyms and words to identify an undergraduate degree as a requirement, and a second list of acronyms and words to identify a professional degree requirement.[34] In Figure 15, we compare the undergraduate attainment/requirements across 4-digit SOC codes, using a local polynomial smoother. Within the individual years that are plotted, the correlations between the two measures are 0.07, 0.65, 0.60, and 0.37. (The correlation for 1990, the unplotted year, is 0.40.)

In Figure 16, we perform the same exercise for professional and post-graduate degrees. Here, the extent to which our data align with educational attainment in the decennial census is substantially weaker. The correlation in the pooled sample of years and occupations is 0.24. For the individual years in our sample, the correlations across SOC codes are 0.15, 0.37, 0.28, 0.30, and 0.16. Overall, we conclude that the educational requirement data which we extract from our newspaper data are correlated with the more cleanly measured data on

[34] These two lists are i) "bachelors," "bachelor," "ba," "bsme," "bs," "bsche," "bsce," "bscs," and "bsee" and ii) "cpa," "masters," "ma," "mba," and "phd."

Figure 13: Occupation Shares

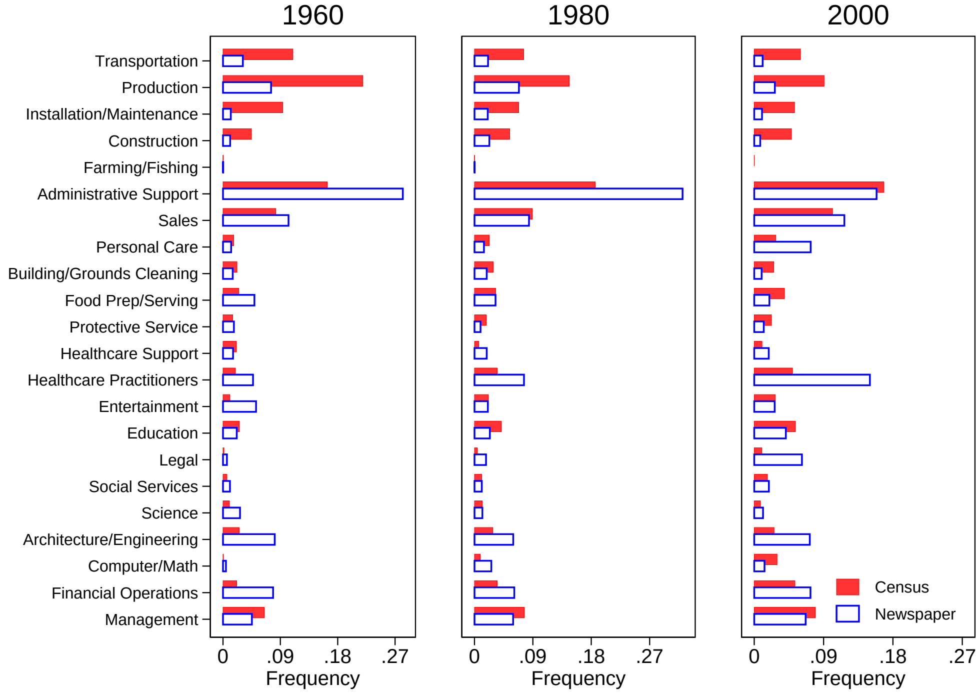


Figure 14: Occupation Shares

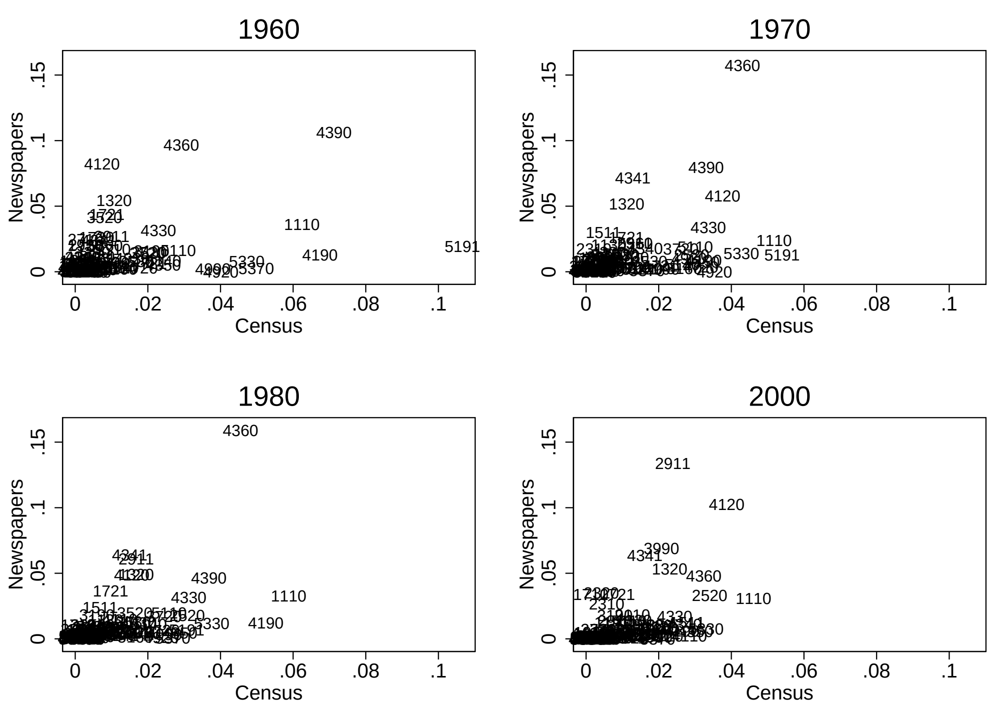
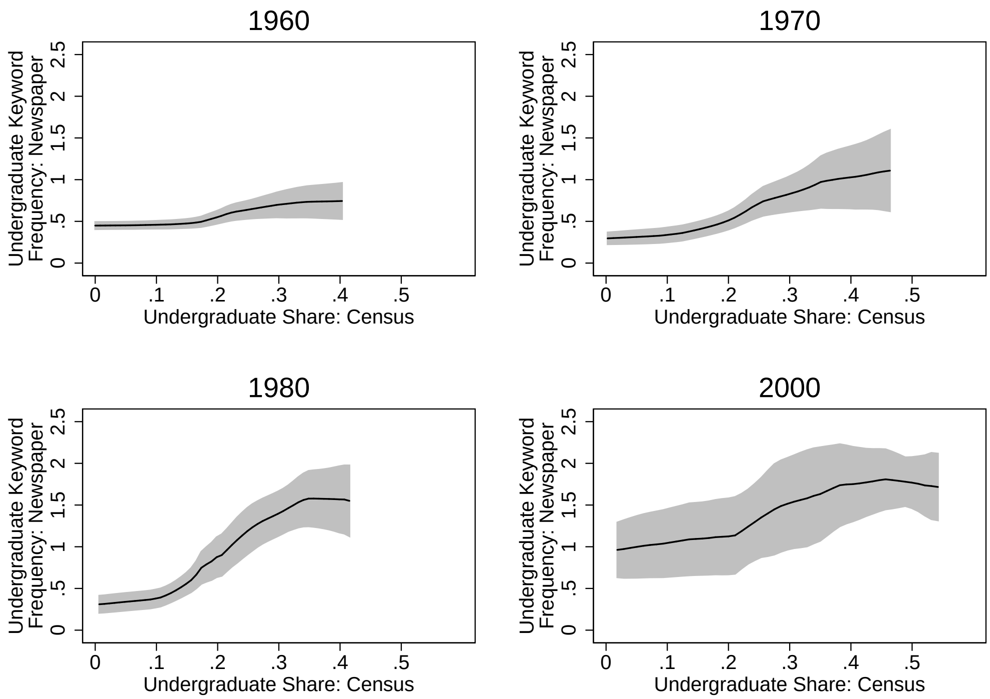


Figure 15: Educational Characteristics by Occupation



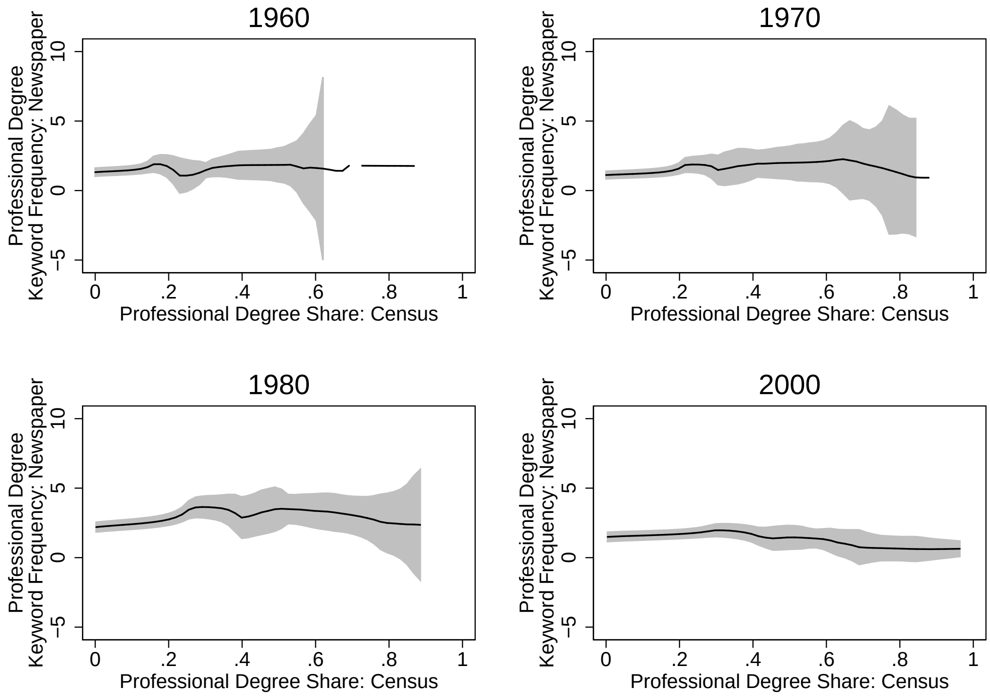
workers' educational attainment, but only moderately correlated for undergraduate degrees, and weakly so for professional and graduate degrees.

B Methods of Job Search

In this section we consider the possibility that the ads which appear in our data set represent a selected sample of all job search methods. In our main analysis, we observe a dramatic increase in words related to nonroutine tasks, which we interpret as reflecting the increasing importance of nonroutine tasks in the economy. But it is also plausible that employers posting vacancies for jobs requiring nonroutine tasks are increasingly likely to post in newspaper ads over time. This section provides empirical evidence that the representativeness of our data set (among the set of all channels of job search) has not changed within the sample period.

Here, we measure the methods that unemployed workers use to search for jobs. We can do this using the IPUMS CPS-ASEC (Ruggles, Genadek, Goeken, Grover, and Sobek, 2015). Unemployed workers are asked whether they have used particular job search methods, and are allowed to report as many methods as they like. The population analyzed here includes

Figure 16: Educational Characteristics by Occupation



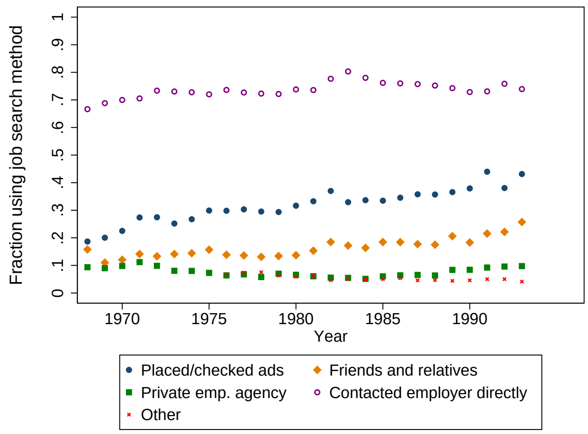
unemployed civilian workers, who are looking for a job, and who are between the ages of 16 and 64. Figure 17 reports the fraction of these workers who use alternative methods for finding a job. The variable of interest for our study is whether the unemployed worker "placed or answered ads" as a method of job search.

Figure 17 shows trends in the method of job search over time. Two methods, placing or answering ads and searching through friends and relatives, increase steadily from 1968 to 1993. Note that by themselves these upward trends are not necessarily problematic; what could pose a problem, however, is the presence of differential trends by occupation, educational background, or other demographic characteristics.

In what follows we consider whether there is selection into job search by task intensity of the worker's 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.

To test this hypothesis, we first compute the mean task content in each occupation over the entire sample period. We then plot the fraction of workers searching for jobs through ads whose last occupation was highly intensive in, say, nonroutine interactive tasks (75th

Figure 17: Methods of Job Search Among Unemployed



Notes: The figure above reports the fraction of unemployed workers who use alternative methods for finding a job. Respondents are allowed to report as many methods as they deem appropriate; therefore the fraction using each method need not sum to 1.

percentile or higher) and the same for workers in occupations that have a low intensity in the same task (25th percentile or lower).[35]

Figure 18 plots the yearly averages for high versus low task intensity occupations. The main takeaway is that while the overall trend is increasing, there does not appear to be a differential trend by the task intensity of the worker's prior occupation. This is reassuring for the main results of the paper because if, for example, the observed rise in interactive tasks were driven by selection of highly interactive job vacancies into newspapers, we would expect workers who work in highly interactive jobs to search more in newspapers over time, relative to workers in low interactive jobs.[36]

We test this hypothesis formally using the following regression:

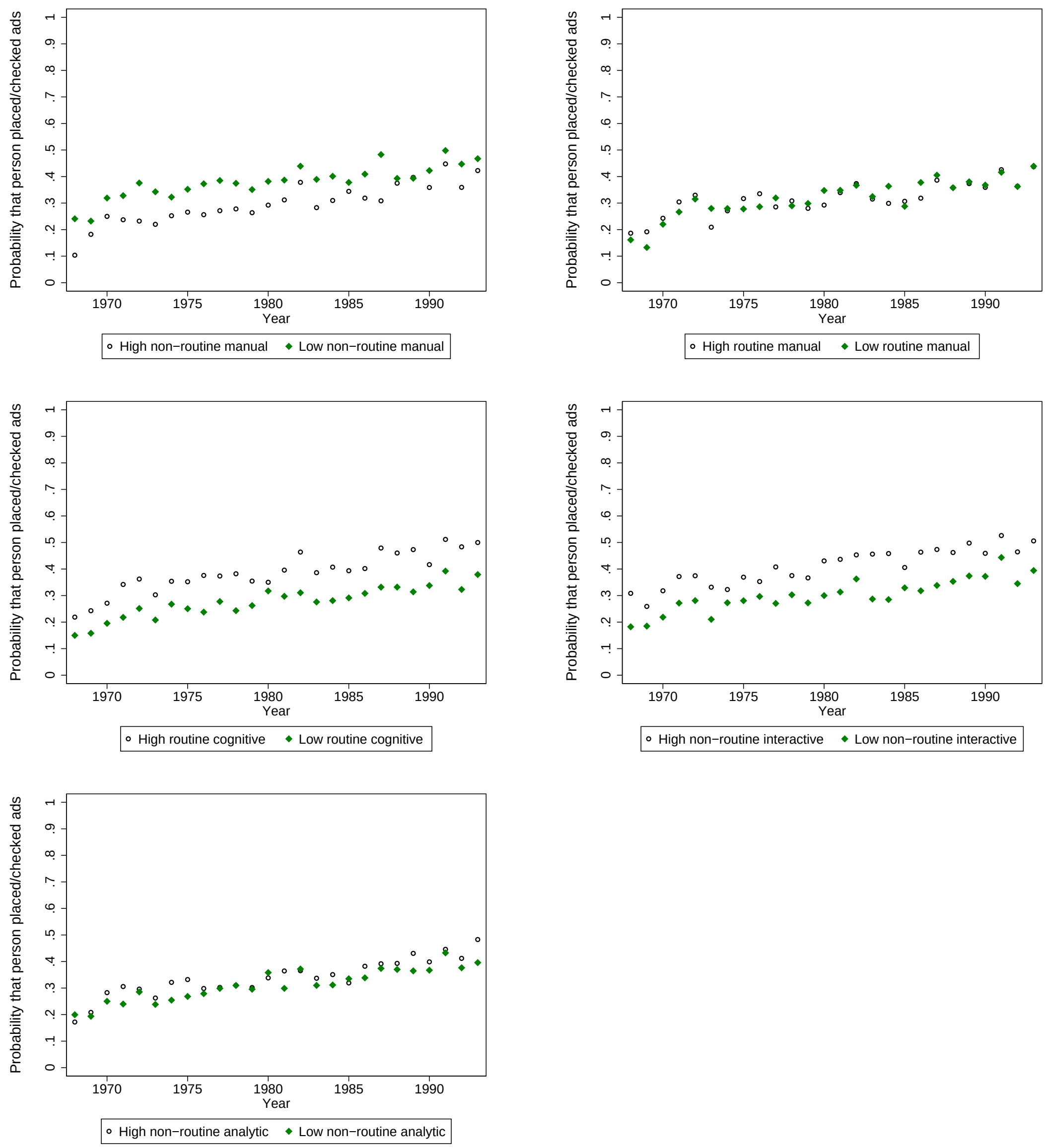
$$y_i = \beta_0 + \beta_t \widetilde{T}_h^\tau + x_i^{'}\gamma + \xi_t + \epsilon_i, \tag{10}$$

where $\widetilde{T}_h^\tau$ is an indicator for being in the τth percentile of the task h distribution (and h indicates one of the five Spitz-Oener task measures). In keeping with our regressions from

[35]The analysis that follows is not sensitive to this choice of threshold for high and low intensity occupations.

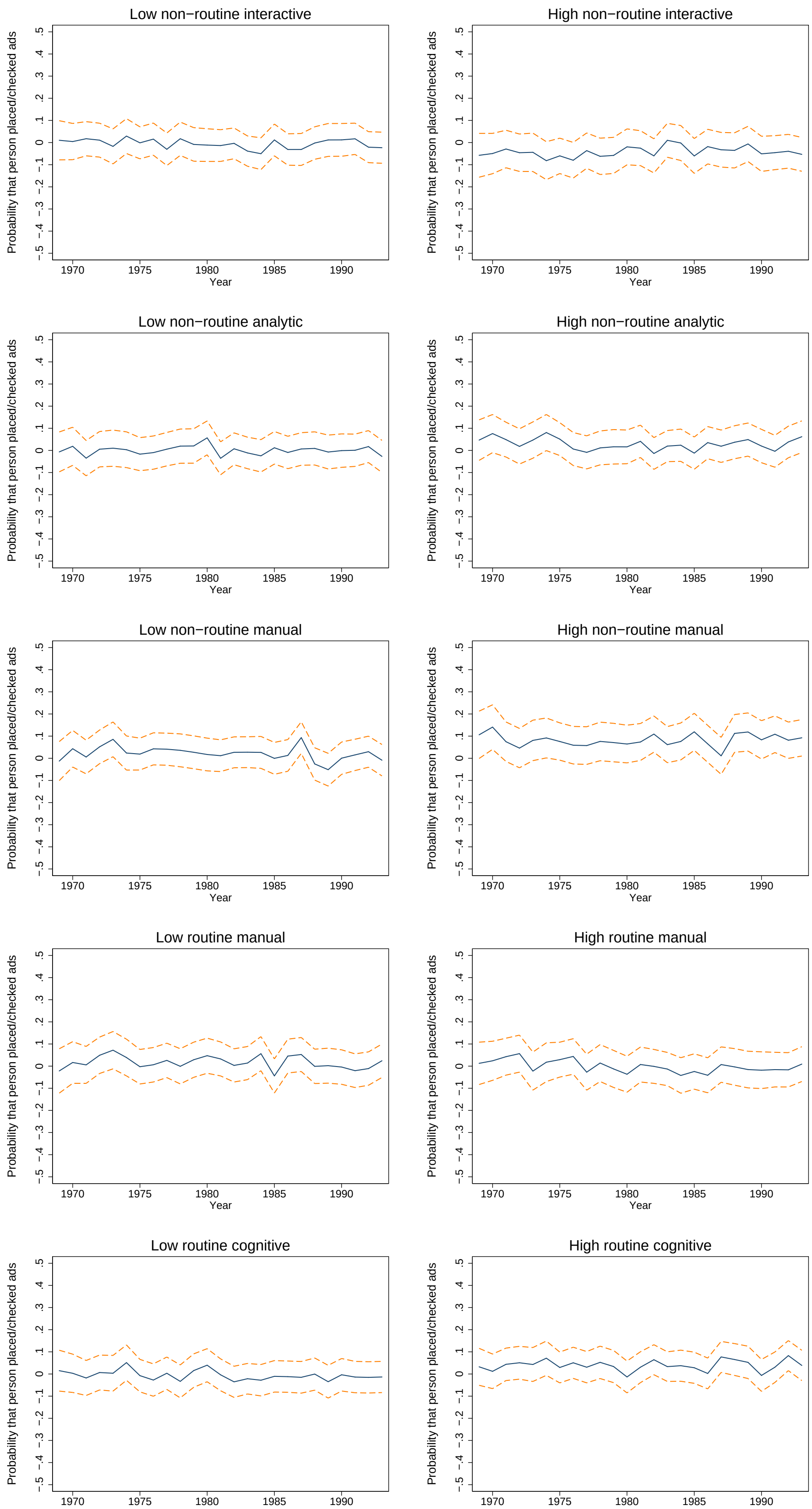
[36]While trends in job search do not appear to differ by occupation, there are *level* differences in job search: For example, workers in high routine cognitive occupations are more likely to use job ads as a search method compared with workers in low routine cognitive occupations. This pattern is consistent with our finding in Appendix A that job vacancies in administrative support occupations are more likely to appear in newspapers when compared to overall employment. That i) occupations with high routine cognitive task content are more likely to appear in newspapers, and ii) workers from these occupations are more likely to search in newspapers is reassuring for this validation exercise, since it demonstrates that demand-side differences in vacancy posting behavior are also reflected in supply side search behavior.

Figure 18: Trends in Job Search Method by Task Intensity of Occupation



Notes: The figure above plots the fraction of job seekers using job ads as a search method, by task intensity of prior occupation. “High” refers to being in an occupation in the 75th percentile or higher in a given task, while “low” refers to being in the 25th or lower percentile in a given task.

Figure 19: Trends in Job Search Method by Task Intensity of Occupation



Notes: The figure above plots the estimates for β_t in Equation 10 for each of the five tasks, along with the 95 percent confidence intervals. Each panel represents the results of a separate regression.

Section 5.1, the vector x_i are controls including gender, marital status, experience dummies (<10 years, 10-19 years, 20-29 years, and 30+ years), a non-white race dummy, and dummy variables over our five educational groups. Figure 19 plots the estimates for β_t. The omitted year is 1968, so coefficients β_t are interpreted as relative to 1968. Overall, Figure 19 suggests no detectable trends in job search behavior through ads.

C Details on the Construction of the Database

This section provides further details that, due to space constraints, we could not include in Section 2. As discussed in that section, constructing the database entails transforming raw, unstructured text into a set of job ads for which we identify job titles and task contents. This requires four steps: i) identifying pages of job ads from the broader sample of advertisements, ii) processing the newspaper text files, iii) grouping occupations according to useful classifications, and iv) eliciting task and skill related information. We turn to each next. Note that some of the language in this appendix is taken directly from Section 2.

C.1 Details on the Latent Dirichlet Allocation Procedure, Used to Distinguish Vacancy Postings from Other Ads

Given the massive amount of newspaper text, it is practically impossible for us to manually distinguish job vacancy postings from other types of advertisements. A simple solution would be to remove newspaper pages where no job vacancy related word can be found. This solution, however, could be problematic as sales-related advertisements, which are a large fraction of advertisements, seem to have several words that appear in job vacancy postings as well. The word "sale" itself would be a perfect example of this. Nevertheless, it is reasonable to assume that job vacancy postings would have different features (distributions of words, to be more precise) compared to other types of advertisements. Using this intuition, we apply Latent Dirichlet Allocation (LDA), an algorithm that categorizes documents within a corpus on the basis of their words. It is an "unsupervised learning algorithm," in the sense that a training data set is not required. The exposition in this section draws heavily from Blei, Ng, and Jordan (2003, pp. 996-998).

Notation and Terminology

1. A *vocabulary* is a set of all possible V words.

2. A *word* w is a vector of length $|V|$. If w takes on the ith element of the vocabulary,

then $w_i = 1$. Otherwise, $w_i = 0$.

3. A *document* is a sequence of N words denoted by $\mathbf{d} = (w_1, w_2 ... w_N)$.

4. A *corpus* is a collection of M documents denoted by $\boldsymbol{D} = (\mathbf{d}_1, \mathbf{d}_2 ..., \mathbf{d}_M)$.

5. A *topic* $z \in \{z_1, z_2, ..., z_K\}$ denotes a "hidden label" across documents in a corpus. The dimensionality K is assumed to be known and fixed.

Data Generating Process

The model assumes the following process.

1. First, choose a vector $\alpha = (\alpha_1, \alpha_2, ..., \alpha_K)$ and a K by V matrix β. Hold these α and β fixed throughout the corpus.

2. Next, for each document $\mathbf{d_m}$ in the corpus, choose a K-dimensional topic weight vector θ_m drawn from a Dirichlet distribution with parameter vector α. That is,

$$\Pr\left(\theta_{m1}, \theta_{m2}, ..., \theta_{mK} | \alpha_1, \alpha_2, ..., \alpha_K\right) = \frac{\Gamma\left(\sum_k \alpha_k\right)}{\prod_k \Gamma\left(\alpha_k\right)} \cdot \prod_{k=1}^{K} \left(\theta_{mk}\right)^{\alpha_k - 1},$$

   where each $\alpha_k > 0$ and $\Gamma\left(\cdot\right)$ refers to the Gamma function.

3. Finally, each word in a document $\mathbf{d}_m$ is determined by first choosing a topic$z \in \{z_1, z_2, ..., z_K\}$ where the probability of choosing a particular topic k is equal to $\Pr(z = z_k | \theta_{m1}, \theta_{m2}, ..., \theta_{mK}) = \theta_{mk}$. Then, choose a word $w_n$from a word-topic probability matrix β where the n, k element of β=$\Pr\left(w_n = 1 | z_k = 1\right)$.

Conditional on α and β, the joint distribution of a topic mixture θ, a set of topics $\mathbf{z}$, and a document $\mathbf{d}_m$ (which contains words w_n) is given by:

$$\Pr\left(\theta, \mathbf{z}, \mathbf{d}_m | \alpha, \beta\right) = p\left(\theta | \alpha\right) \prod_{n=1}^{N} p\left(z_k | \theta\right) p\left(w_n | z_k, \beta\right).$$

The marginal distribution, or likelihood, of a document $\mathbf{d}_m$ which contains words w_n is given by integrating over θ and summing over potential topics z_k:

$$\begin{aligned}
\Pr\left(\mathbf{d}_m | \alpha, \beta\right) &= \int p\left(\theta | \alpha\right) \left(\prod_{n=1}^{N} \sum_k p\left(z_k | \theta\right) p\left(w_n | z_k, \beta\right)\right) d\theta \\
&= \frac{\Gamma\left(\sum_k \alpha_k\right)}{\prod_k \Gamma\left(\alpha_k\right)} \cdot \int \prod_{k=1}^{K} \left(\theta_{mk}\right)^{\alpha_k - 1} \prod_{n=1}^{N} \sum_k \prod_{v=1}^{V} \left(\theta_{mk} \beta_{kv}\right)^{w_n} d\theta.
\end{aligned}$$

Estimation

The main purpose of LDA is to determine the distribution of the latent topics conditional on the observed words in each document. The distribution is as follows:

$$\Pr\left(\theta, \mathbf{z}|\mathbf{d}_m \alpha, \beta\right) = \frac{\Pr\left(\theta, \mathbf{z}, \mathbf{d}_m|\alpha, \beta\right)}{\Pr\left(\mathbf{d}_m \alpha, \beta\right)}.$$

The estimated values, $\widehat{\alpha}$ and $\widehat{\beta}$, are values of α and β which maximize the log likelihood of the documents:

$$\left(\widehat{\alpha}, \widehat{\beta}\right) = \arg\max \left\{ \sum_{m=1}^{M} \log\left[\Pr\left(\mathbf{d}_m|, \alpha, \beta\right)\right] \right\}.$$

The posterior distribution cannot be computed directly. As a feasible approximation, we use Hoffman, Blei, and Bach (2010)'s Expectation Maximization algorithm. The python code for this algorithm is part of the *gensim* module; see Rehurek and Sojka (2010).

Details on Our Implementation

To construct our LDA model, we take samples of pages of advertisements for each of our newspapers, separately: display ads in the *Boston Globe*, spanning 1960 to 1983; display ads in the *New York Times*, 1940 to 2000; classified ads in the *Boston Globe*, 1960 to 1983; classified ads in the *New York Times*, 1940 to 2000; and classified ads in the *Wall Street Journal*, 1940 to 1998. Since the text in display ads is larger, our code more easily identifies and processes the text in these ads. For this reason, we apply our processing code separately for these different types of ads.

For each of our five subsamples, we first restrict attention to pages of advertisements which are at least 200 words. From these pages of ads, we remove *stop words* (e.g., common words like "a," "the," and "and"), numerals, and words that are not contained in the English dictionary. We then *stem* words; that is, we remove word affixes so that words in different forms—singular nouns, plural nouns, verbs, adjectives, and adverbs—are grouped as one. To emphasize, the removal of certain types of words and the stemming of words pertains to the construction of our LDA model. Once we have estimated this model, we will restore our original text.

After estimating the model, each page in each subsample is defined by a probability distribution over K topics. We pick the pages for which the probability of belonging to the topic with words common in vacancy postings is greater than 0.40. The choice of the cutoff balances a trade-off between throwing out too many vacancy postings (particularly job ads in sales-related occupations) and including too many non-job ads in our data set. Choosing a low cut-off will lead us to include pages of non-job-related ads at this stage. However, since succeeding stages will discard ads without job titles, these pages of non-job ads will be excluded.

LDA Results

In Table 1, using the subsample of text from the *Boston Globe* display ads, we presented partial results from our LDA procedure, describing the five words which were most predictive of the five different topics. In Table 7, we provide analogous results for the all five subsamples: the *Boston Globe* classified ads, the *Boston Globe* display ads, the *New York Times* classified and display ads, and the *Wall Street Journal* classified ads. Again, we chose the number of topics, K, so that i) there is a clear, identifiable topic associated with job ads, but ii) if we were to add a $K+1^{\text{st}}$ topic, then there would be multiple job-related topics resulting from the LDA model. The words presented in these tables are those with the highest values of β_{nk}.

C.2 Processing the Newspaper Text Files

The goal of this step is to parse the raw text files to produce a set of separate ads, complete with job titles and word content.

Discerning the Boundaries between Vacancy Postings In the ProQuest data set, the advertisements on a single page are all grouped in a single text field. Our next task is to identify when one vacancy posting ends and a second posting begins. Here, certain phrases at the beginning or end of individual help wanted ads allow us to identify the boundaries between ads. We use the following three-step rule to demarcate individual ads:

- Addresses: Most job vacancy posts have the employers' addresses at the end. The first step of our algorithm marks the end of an advertisement. We are able to match zip codes, such as "Boston MA 02107," and street addresses, such as "82 DEVONSHIRE STREET."

- Ending phrases: Some phrases indicate that a job vacancy post is ending, for example: "send [...] resume," "submit [...] resume," "in confidence to," "affirmative action employer," "equal opportunity;" or "equal opportunities." The algorithm marks the end of an advertisement if any of these patterns is detected and starts a new one after the following line.

Table 7: Predictive Word Stems in the LDA Model

Panel A: Boston Globe Classified	
1	opportun experi work call salari year employ posit manag resum
2	offic new day avail want free inc busi servic mass
3	auto new car call ford low stereo rte price motor
4	bdrm bath kit mod new bay condo home area back
5	bdrm bay mod back avail kit bath studio call build
Panel B: Boston Globe Display	
1	system experi comput opportun year manag engin program requir design
2	reg save store size price color style set regular charg
3	car price new auto power tire stereo air door stock
4	day free new one call coupon travel per offer week
5	street open inc ave mass call rte rout new home
Panel C: New York Times Classified	
1	resum seek call must work exp excel new salari send
2	new home owner acr call car hous den area ask
3	build ave new park call studio east avail fee fir
Panel D: New York Times Display	
1	experi manag system comput new opportun program requir year salari
2	school sat sun new call music ticket program wed art
3	day hotel travel free night includ new call per beach
4	one get make new year help take like time busi
5	new room call home ave avail park build floor offic
6	new white size avenu black store fifth floor wool open
7	free call new order phone charg card pleas send mail
8	rate new fund bank invest offer inc may compani interest
9	ave new street avenu park plaza unit east mall twin
10	car new call auto leas drive mile dealer power price
11	new book time world art film one magazin page news
12	price reg save design color set furnitur store select rug
Panel E: Wall Street Journal Classified	
1	experi resum salari opportun posit year market requir send develop
2	acr home call room new view beach pool bath hous
3	estat real properti offic leas locat unit call build new
4	busi street journal box wall call compani servic new avail
5	new call air owner price interior avail number ad car

- Beginning of the posts: A job vacancy post usually starts with a job title, which stands alone within a few lines and is uppercase. If we detect consecutive short lines of uppercase letters, we group those lines together. We then test whether the lines are a job title (see below). If so, the algorithm assigns the beginning of an advertisement here.

Identifying the Advertisement's Job Title Finally, since one of the main goals of the project is to identify how individual occupations' skill profiles have changed over time, it will be necessary to assign each vacancy posting to its corresponding occupation. On their website, O*NET publishes, for each occupation, a "Sample of Reported Job Titles." We retrieve these titles, a list of more than eight thousand, from the O*NET website. From this list of titles, we construct a list of one-word personal nouns. For instance, "Billing, Cost, and Rate Clerks" is a potential job title according to O*NET; see http://www.onetonline.org/link/summary/43-3021.02. Since it is exceedingly unlikely that this exact phrase appears in any of our ads, we identify a potential job title to appear when the word "Clerk" is mentioned.

Then, in each line in which either i) words are all-capitalized or ii) only one or two words appear, we search among the words in that line of advertisement text for a personal-noun job title. According to our example in Figure 1, this would occur in the lines containing "MUTUAL FUNDS CLERKS," "DEVONSHIRE STREET BOSTON MA 02109," "PAYMENTS CLERKS," "TERMINAL OPERATOR," "Fidelit," "Group," and "111." The first four examples would satisfy criterion (i). The lines with "MUTUAL FUNDS CLERKS," "PAYMENTS CLERKS," and "TERMINAL OPERATOR," also contain a personal-noun job title. As a result, the contents of these lines are reported as the ads' job titles.

C.3 Details on the Continuous Bag of Words Model

The goal of the continuous bag of words (CBOW) model is to compute the similarity among words or phrases in our corpus. The first three subsections of this appendix provide a basic set up of such a model, drawing from Section 2 of Bengio, Ducharme, Vincent, and Jauvin (2003). After this, we describe how we use our estimated continuous bag of words model to link job titles to SOC codes and job ad text to categories of occupational characteristics.

Notation and Terminology

1. A *word,* w_i, is a unit of text.

2. A *vocabulary,* V, is a set of all possible words.

3. A *corpus* is a sequence of words denoted by $\{w_1, w_2 ... w_T\}$.

4. A *context* of a word is a set of adjacent words of predetermined distance. For example, in the n-gram model, a context of a word w_i is $\mathcal{H} = \{w_{i-1}, w_{i-2}, ... w_{i-n}, w_{i+1}, ..., w_{i+n}\}$, a set of $2n - 1$ words which appear within a n-word window of w_i .

Model Setup and Estimation

The underlying assumption in the continuous bag of word model is that words in similar contexts share semantic meaning in the population of text data. In the CBOW model, similar context refers to a set adjacent words, typically a fixed number of n words surrounding the word..[37]

The objective, which we will try to maximize via maximum likelihood, is given by the probability of observing a word w_t conditional on the features of the words in its context $C(\mathcal{H})$. Below, we will use $\hat{P}(w_i|C(\mathcal{H}))$ to denote this probability (which is our MLE objective). The model estimation can be divided into two parts:

1. A mapping C from each word w_i in V to a real vector of predetermined length N. Here N will be parameter we are free to choose, describing how many features to include in our model to describe each individual word. In practice, C is a |V| by N dimensional matrix.

2. A function g which maps a sequence of C words in the context to a conditional probability distribution over words. $\hat{P}(w_i|C(\mathcal{H})) = g(w_i, C(w_{j1}), C(w_{j2}), ...)$, where all of the j belong to $\mathcal{H}$. In practice, g can be represented as a N by |V| matrix for each possible context $\mathcal{H}$.

In these two steps, we are predicting the likelihood of observing a particular word w_i based on the features of the words that surround it. Our model will yield a good representation of the words in our vocabulary if they accurately and parsimoniously describe the attributes of these words. The maximum likelihood procedure chooses C and g to match conditional probabilities observed in the corpus. Though the idea is relatively simple, the dimensionality of the model requires additional adjustments to reduce the computational burden. To this end, we follow the procedure as mentioned in Mikolov et al. (2013a; 2013b). We choose the dimension of C to be 300, and the context of word w_i to include the five words succeeding and preceding each w_i. In writing out of MLE objective function, we omit words w_i which appear fewer than five times in our corpus.[38]

[37] However, given the same set of adjacent words, the order does not matter. For example, if the window size is 1, then, the context [..., w_1, $\overline{w}$, w_2, ...] is the same as [..., w_2, $\overline{w}$, w_1, ...] for any word $\overline{w}$.

[38] Mikolov, Chen, Corrado, and Dean (2013a) estimated a model using text from a Google News corpus,

Construction of the CBOW Model

As part of a separate ongoing project, EMSI has provided us a wide sample of job ads posted online between April 2011 and March 2017. As with our newspaper data, these text contain a job title for each vacancy posting. Unlike the newspaper data, and useful for the construction of a mapping between job titles and SOC codes, the EMSI data include a EMSI-constructed SOC code for each advertisement.

Using our sample of text from the *Boston Globe*, *New York Times*, and *Wall Street Journal* and the entries from sample of 4.2 million job ads that were posted in January 2012 or January 2016, we construct a continuous bag of words model, applying the procedure outlined in the previous subsection. The output of this model is a vector representation, C, for each word. A phrase, too, can be represented as a vector, as the sum of the vectors of the phrase's constituent words. For example, we will find it useful to construct a vector representation of a phrase like "construction manager." To do so, we would simply sum the vectors for "construction" and "manager."

With our estimate of C we use a cosine similarity score: $\frac{C(w_i)\cdot C(w_j)}{|C(w_i)||C(w_j)|}$ to compute similarity between two words (or phrases) w_i and w_j. We use this similarity score for two purposes: to link job titles to SOC codes, and to link the words used in the body of job ads to categories of work characteristics. In the following two subsections, we detail these two applications of our CBOW model.

Grouping Occupations and Mapping Them to SOC Codes

In the newspaper data, postings for the same occupation appear via multiple distinct job titles. For example, vacancy postings for registered nurses will be advertised using job titles which include "rn," "registered nurse," or "staff nurse." These job titles all map to the same occupation: 291141 using the BLS Standard Occupational Classification (SOC) system, or 3130 according to the 2000 to 2009 vintage of the Census Occupation Code. To group job titles to occupation codes, we apply the BLS SOC code. We first lightly edit job titles to reduce the number of unique titles: We combine titles which are very similar to one another (e.g., replacing "host/hostesses" with "host," and "accounting" with "accountant," etc.); replace plural person nouns with their singular form (e.g., replacing "nurses" with "nurse," "foremen" with "foreman," etc.); and remove abbreviations (e.g., replacing "sr" with "senior," "asst" with "assistant," and "customer service rep" with "customer service representative").

and found that increasing the dimensionality of the model's vector representation from 300 to 600 led to only small improvements in performance.

From this shorter list, we apply two complementary methods to classify ads according to SOC codes. First, for the most common 1000 job titles from the *Boston Globe*, *New York Times*, and *Wall Street Journal* job ads, we apply the mapping of www.onetsocautocoder.com to retrieve the SOC codes. These top 1000 job titles span 2.2 of the 6.6 million ads from the previous subsection's processed newspaper data.

For the remaining 4.4 million ads, we apply a *continuous bag of words* model in combination with an ancillary data set provided to us by EMSI (see Bengio, Ducharme, Vincent, and Jauvin, 2003, and Mikolov et al., 2013a; 2013b). Generally speaking, a continuous bag of words 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 "nurse" and "rn" both tend to appear next to words like "patient," "medical," or "acute" one would conclude that "nurse" and "rn" have similar meanings to one another. Building on this idea, a continuous bag of words model represents each word as a (long) vector, with the elements in each vector measuring the frequency with which other words are mentioned nearby (e.g., for the "nurse" vector, what fraction of the time in our corpus of vacancy posting text are "aardvark, "abacus," ... "zoo," or "zygote" mentioned in close proximity to the word "nurse"?). Given this vector representation, two words are similar if the inner product of their vectors is large. Short phrases, too, can be usefully represented as vectors as the sum of the vectors of the constituent words (for example, the vector representation of "construction manager" would equal the sum of the "construction" and "manager" vectors.) Taking stock, with the continuous bag of words model, we can represent any phrase—and, in particular, any job title—as a vector. As a result, we can also compute the similarity of any two job titles, as the inner product of the job titles' associated vectors.

This continuous bag of words model is useful, since we have an ancillary data set containing a large correspondence between job titles and SOC codes. With this model in hand, for a given job title in our newspaper text we can compute the closest job title in the EMSI data set, and therefore retrieve the SOC code for that job title. We perform this second procedure on all of the job titles which appear in at least two ads in our data set. In combination with the first approach, we have been able to collect SOC codes for 4.1 million job ads.

We believe approach the based on www.onetsocautocoder.com is more accurate. However, it is infeasible to hand-collect the hundreds of thousands of possible job titles which appear in our newspaper text. For the 1,000 most common job titles, we can examine how our continuous bag of words procedure matches with the hand-collected SOC codes. The SOC codes resulting from the two approaches agree at the 2-digit level 80 percent of the time, at the 4-digit level 67 percent of the time, and at the 6-digit level 59 percent of the time.

Eliciting Skill- and Task-Related Information

Within the body of the job ads, similar words will refer 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. Here, we construct three classification schemes.

Our main classification follows that of Spitz-Oener (2006) who, in her study of the changing task content of German occupations, groups activities into five categories: *nonroutine analytic, nonroutine interactive, nonroutine manual, routine cognitive,* and *routine analytic.* In our main application of her categorizations, we begin with the words in each of her five lists of task-related words. For each list, we append words which are similar to those in footnote 14, where similarity is determined by our continuous bag of words model: We append words which have a cosine similarity greater than 0.55 to any of the words in footnote 14. We also append any additional words which have one of the ten highest cosine similarity scores in each task group. This is our primary classification, and we use it in each calculation that follows in the paper. In addition, as a robustness check, we will consider a narrower mapping between categories and words, one which only relies in Spitz-Oener (2006)'s definitions as enumerated in footnote 14.

For varying purposes, we also consider two additional complementary classifications. With the aim of emulating O*NET's database, we construct our own mappings between words and phrases on the one hand and occupational work styles (corresponding to O*NET Elements beginning with 1C), skills (encompassing O*NET Elements 2A and 2B), knowledge requirements (corresponding to O*NET Elements 2C), and work activities (O*NET Elements 4A) on the other. For each O*NET Element, we begin by looking for words and phrases related to the O*NET Title and, refer to the O*NET Element Description to judge whether these synonyms should be included, as well as if other words should be included. For instance, for the "Production and Processing" knowledge requirement, our list of synonymous words includes the original "production" and "processing," as well as "process," "handle," "produce," "render," and "assembly." And since the O*NET Description for "Production and Processing" states that the skill is associated with the "Knowledge of raw materials, production processes, quality control, costs, and other techniques for maximizing the effective manufacture and distribution of goods," we also include "quality control," "raw material," "qc," and "distribution" in our list of words and phrases to search for when measuring this knowledge requirement. Admittedly, since this procedure is based on our own judgment, it is necessarily ad hoc. Moreover it will not be able to capture all of the words phrases which are indicative of a particular work style, skill, knowledge requirement, or work activity.

For this reason, we append to our initial lists of words and phrases an additional set of words, using a continuous bag of words model similar to the one constructed in Section 2.2, built from the newspaper (1960 to 2000) and online (January 2012 and January 2016) job ads. We compute the similarity of the words in each O*NET Element Title and all of the other words in our corpus of newspaper and online vacancy postings. For instance, for "Production and Processing," our model yields: "process," "processes," "packaging," "preparation," and "manufacturing" as the words with the highest cosine similarity. We take the top 10 words, plus any additional words which have a cosine similarity greater than 0.45, to the O*NET Element Title and add these words to those words and phrases from our "judgment-based" procedure described in the previous paragraph.

Each of the two approaches, the "judgment based" procedure and the "continuous bag of words model based" procedure, has its strengths and weaknesses. On the one hand, the first procedure is clearly ad hoc. Moreover, the continuous bag of words model has the advantage of accounting for the possibility that employers' word choice may differ within the sample period.[39] On the other hand, there is a danger that the continuous bag of words model will identify words to be similar even if they are not synonymous. For example, the vector representations in our bag of words model indicates that the five most similar words to the "Mathematics" O*NET Element Title are "math," "physics," "economics," "algebra," and "science." While the first five words strike us as reasonable, a word like "linguistics" (which also appears in the list of similar words according to the CBOW model) does not. In our job ads, since (apparently) mathematics and linguistics are mentioned in similar contexts, our model suggests that these words have similar meaning when in fact they do not.

Our third classification scheme applies the mapping between keywords and skills which Deming and Kahn (2017) define in their study of the relationship between firms' characteristics and the skill requirements in their vacancy postings.[40] To each of these lists of words, we

[39]For instance, even though "creative" and "innovative" largely refer to the same occupational skill, it is possible that their relative use among potential employers may differ within the sample period. This is indeed the case: Use of the word "innovative" has increased more quickly than "creative" over the sample period. To the extent that our ad hoc classification included only one of these two words, we would be mis-characterizing trends in the O*NET skill of "Thinking Creatively." The advantage of the continuous bag of words model is that it will identify that "creative" and "innovative" mean the same thing because they appear in similar contexts within job ads. So, even if employers start using "innovative" as opposed to "creative" part way through our sample, we will be able to consistently measure trends in "Thinking Creatively" throughout the entire period.

[40]See Table 1 of Deming and Kahn (2017) for their list of words and their associated skills. Building on their definitions, we use the following rules 1) cognitive: analytical, cognitive, critical thinking, math, problem solving, research, statistics; 2) social: collaboration, communication, negotiation, presentation, social, teamwork; 3) character: character, energetic, detail oriented, meeting deadlines, multi-tasking, time management; 4) writing: writing; 5) customer service: client, customer, customer service, patient, sales; 6) project management: project management; 7) people management: leadership, mentoring, people manage-

append additional words which are sufficiently similar (those with a cosine similarity greater than 0.55 or among the 10 most similar words for each category) to any of the words in the original list.

D Sensitivity Analysis Related to Sections 3 and 4

In this appendix, we explore the results of Sections 3 and 4 with alternate skill and task measures and alternate occupation classification schemes.

D.1 Top Occupations

In addition to the measures developed by Spitz-Oener (2006) and Firpo, Fortin, and Lemieux (2014) that we explored in the body of the paper, we apply the mapping between keywords and skills that Deming and Kahn (2017) used in their study of the relationship between firms' characteristics and the skill requirements in their vacancy postings. For each skill group, we append words that are similar to those mentioned in footnote 40, using the continuous bag of words model to identify words and phrases that are similar to one another.

Table 8 lists the top occupations according to the skill groups in Deming and Kahn (2017). According to intuition, ads for sales representatives and sales managers have the highest frequency of words related to customer service skills; ads for financial specialists and managers have the highest frequency of words related to financial skills; and ads for supervisors, managers, and health diagnosticians have the highest frequency of words related to people management.

D.2 Trends in Keyword Frequencies

In Section 4.1, we considered trends in the frequency with which different groups of words were mentioned in our newspaper text. We showed that the frequency of words related to routine cognitive and routine manual tasks have declined over the sample period, while words related to nonroutine task words have increased in frequency. Moreover, nearly all of these changes have occurred within, rather than between, 4-digit SOC codes. In this appendix, we first consider alternate groups of skill- and task-related words, the first due to Deming and Kahn (2017) and the second due to Firpo, Fortin, and Lemieux (2014). Next, returning to Spitz-Oener (2006)'s occupation codes, we re-consider trends in task mentions

ment, staff, supervisory; 8) financial: accounting, budgeting, cost, finance, financial; 9) computer (general): computer, software, spreadsheets.

Table 8: Top Occupations According to the Deming and Kahn (2017) Classification of Skills

Character			Computer		
1120: Sales Managers	70.1	0.51	1511: Computer	185.4	0.70
1110: Top Executives	127.7	0.48	5350: Water Transportation	0.5	0.45
4140: Sales Rep., Whole./Man.	64.1	0.41	1720: Engineers	69.3	0.43
1191: Other Management	60.3	0.41	1721: Engineers	132.1	0.33
1130: Operation Managers	101.8	0.41	5120: Assemblers/Fabricators	9.8	0.33
Customer Service			**Financial**		
4140: Sales Rep., Whole./Man.	64.1	0.72	1320: Financial Specialists	194.6	0.85
1120: Sales Managers	70.1	0.64	1130: Operation Managers	101.8	0.70
4120: Retail Sales	172	0.54	1311: Business Operations	71.3	0.34
4130: Sales Rep., Services	100.4	0.50	1930: Social Scientists	5.2	0.31
4190: Other Sales	42.8	0.50	1110: Top Executives	127.7	0.28
People Management			**Problem Solving**		
1131: H.R. Managers	21.2	0.52	1520: Math. Science	10.7	0.49
1110: Top Executives	127.7	0.39	1930: Social Scientists	5.2	0.44
1191: Other Management	60.3	0.38	1940: Science Technicians	11.4	0.42
1130: Operation Managers	101.8	0.35	1920: Physical Scientists	20.4	0.41
3710: Cleaning Supervisors	5.2	0.33	1910: Life Scientists	9.0	0.38
Project Management			**Social**		
1721: Engineers	132.1	1.03	2110: Social Workers	50.9	0.15
1720: Engineers	69.3	1.01	1311: Business Operations	71.3	0.11
1511: Computer	185.4	0.67	1131: H.R. Managers	21.2	0.11
1710: Architects	60.3	0.62	2510: Postsecondary Teachers	8.1	0.11
1311: Business Operations	71.3	0.54	1511: Computer	185.4	0.10
Writing					
2730: Media	75.5	0.35			
1930: Social Scientists	5.2	0.12			
1311: Business Operations	71.3	0.07			
2530: Instructors	3.3	0.07			
2990: Other Healthcare	0.7	0.06			

Notes: This table lists the top five occupations according to the frequency with which different activity-related words are mentioned. Within each panel, the first column gives the SOC code and title; the second column gives the number of job ads in our data set (in thousands); and the final column gives the frequency (mentions per vacancy posting) of task-related words. Only occupations with at least 200 advertisements, representing 104 4-digit SOC codes, are included. Footnote 40 contains the words and phrases corresponding to each skill that were used in Deming and Kahn (2017). To these lists, we append similar words, using the continuous bag of words model introduced in Section 2.3.

Table 9: Trends in Keyword Frequencies

	Frequency								Within
	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	Share
Character	4.72	5.13	5.41	6.20	6.84	7.14	6.79	5.95	0.90
	[4.71]	[5.09]	[5.31]	[6.29]	[6.82]	[7.06]	[7.02]	[5.81]	(0.08)
Computer	1.05	1.29	1.38	1.90	2.23	2.76	3.80	4.50	1.09
	[1.05]	[1.30]	[1.38]	[1.89]	[2.15]	[2.83]	[4.00]	[4.82]	(0.03)
Customer Service	2.90	2.73	2.99	3.55	3.76	4.69	5.16	5.21	0.90
	[2.89]	[2.67]	[2.89]	[3.53]	[3.72]	[4.51]	[4.95]	[4.97]	(0.04)
Financial	1.95	2.02	1.98	1.90	2.17	2.31	2.44	2.60	0.69
	[1.92]	[1.94]	[1.84]	[1.73]	[1.97]	[2.04]	[2.22]	[2.38]	(0.12)
People Management	2.24	2.77	2.91	3.11	3.36	3.44	3.34	3.06	0.89
	[2.24]	[2.75]	[2.89]	[3.06]	[3.25]	[3.22]	[3.19]	[2.97]	(0.08)
Problem Solving	1.31	1.29	1.06	1.19	1.30	1.52	1.81	2.01	0.73
	[1.30]	[1.26]	[1.02]	[1.10]	[1.18]	[1.35]	[1.74]	[1.81]	(0.06)
Project Management	2.88	2.95	2.91	3.42	3.53	3.61	3.94	4.69	0.91
	[2.89]	[2.97]	[2.95]	[3.42]	[3.35]	[3.38]	[3.70]	[4.54]	(0.03)
Social	0.39	0.45	0.49	0.64	0.95	1.41	1.67	1.96	0.91
	[0.38]	[0.44]	[0.48]	[0.60]	[0.89]	[1.30]	[1.63]	[1.81]	(0.03)
Writing	0.47	0.42	0.39	0.47	0.52	0.71	0.87	0.80	0.90
	[0.46]	[0.40]	[0.35]	[0.39]	[0.46]	[0.66]	[0.77]	[0.76]	(0.05)

Notes: See the notes for Table 5.

with an alternate methodology for identifying task-related words and alternate occupation classifications.

In Table 9, we use Deming and Kahn's categorization of skills. Computer skills, Social skills, and Project Management have become increasingly prevalent with within-occupational changes, again accounting for an overwhelming proportion of these trends.

Next, we study within and between changes in work frequencies using the Firpo, Fortin, and Lemieux (2014) categorization of O*NET work activities and work contexts.[41] Table 10

[41] According to the Firpo, Fortin, and Lemieux (2014) categorization, the O*NET Elements associated with Information Content include codes related to Getting Information (4.A.1.a.1), Processing Information (4.A.1.a.2), Analyzing Data (4.A.2.a.4), Interacting with Computers (4.A.3.b.1), and Documenting/Recording Information (4.A.3.b.6). The elements associated with the Automation/Routine group include Degree of Automation (4.C.3.b.2), Repeating Same Tasks (4.C.3.b.7), Structured versus Unstructured Work (4.C.3.b.8), Pace (4.C.3.d.3), and Repetitive Motions (4.C.2.d.1.i). Face-to-Face Contact maps to Face-to-Face Discussions (4.C.1.a.2.l), Establishing Interpersonal Relationships (4.A.4.a.4), Caring for Others (4.A.4.a.5), Working Directly for the Public (4.A.4.a.8), and Coaching and Developing Others (4.A.4.b.5). The On-site-Job group of elements includes Inspecting Equipment, Structures, or Material (4.A.1.b.2), Handling and Moving Objects (4.A.3.a.2), Controlling Machines and Processes (4.A.3.a.3), Operating Vehicles (4.A.3.a.4), Repairing Mechanical Equipment (4.A.3.b.4), and Repairing Electronic Equipment (4.A.3.b.5). Finally, Decision-Making maps to the following O*NET Elements: Making Decisions (4.A.2.b.1), Thinking Creatively (4.A.2.b.2), Developing Objectives (4.A.2.b.4), Responsibility for Outcomes and Results (4.C.1.c.2), and Frequency of Decision Making (4.C.3.a.2.b).

Table 10: Trends in Keyword Frequencies

	Frequency								Within
	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	Share
Information	2.99	3.32	3.44	4.11	4.54	5.10	6.09	7.03	1.03
Content	[2.97]	[3.32]	[3.41]	[4.02]	[4.33]	[5.08]	[6.27]	[7.14]	(0.03)
Face-to-	7.09	7.55	8.21	9.64	9.82	9.04	9.61	11.21	0.85
Face	[7.08]	[7.63]	[8.21]	[9.52]	[9.53]	[8.46]	[8.97]	[10.60]	(0.03)
On-Site	14.48	13.54	12.65	14.56	11.81	8.36	7.77	7.60	0.86
Job	[14.53]	[13.81]	[12.93]	[15.15]	[12.27]	[9.06]	[8.36]	[8.62]	(0.03)
Decision-	2.00	2.27	2.33	2.71	2.85	2.87	3.54	4.53	0.87
Making	[2.00]	[2.26]	[2.30]	[2.63]	[2.69]	[2.65]	[3.23]	[4.20]	(0.03)

Notes: See the notes for Table 5.

illustrates that words related to Information Content, Face-to-Face Contact, and Decision-Making have increased in prevalence, and that most of these changes are due to within, rather than between, occupational changes.

In the remaining tables of this subsection, we revert to the Spitz-Oener (2006) categorization of groups of tasks. When we produced Table 5, we applied not only Spitz-Oener (2006) mapping between words and task groups, but also our own continuous bag of words model to identify additional words to search for. In Table 11, we recompute trends in keyword frequencies, now using Spitz-Oener (2006)'s original mapping between words and task groups. In this table, the frequency of task-related keywords is lower than in Table 5. However, the trends in keyword frequencies — increasing for nonroutine tasks and decreasing for routine tasks— are similar to those depicted in Table 5. Moreover, as in Table 5, a large fraction of the overall changes in keyword frequencies occur within occupations.

In Table 5, in the body of the paper, we applied a 4-digit SOC classification. In Tables 12 and 13, we recompute these trends with narrower occupation classification schemes. In Table 12, we apply a 6-digit SOC classification, while in Table 13 we categorize ads according to their job titles (as opposed to grouping job titles by SOC codes). The classification used in Table 13 is the finest classification one could possibly apply when decomposing trends in keyword frequencies in between-occupation and within-occupation components. As one would expect, with a finer occupation classification scheme, the share of changes in keyword frequencies that are due to the within-occupation component are somewhat lower. However, even here, for each of the task measures, a majority of the changes in task content occur within occupations.

Table 11: Trends in Keyword Frequencies

	Frequency								Within
	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	Share
Nonroutine	1.92	2.08	1.95	1.94	2.03	2.10	2.27	2.62	0.82
Analytic	[1.91]	[2.08]	[1.93]	[1.91]	[1.95]	[1.92]	[2.15]	[2.49]	(0.11)
Nonroutine	1.74	1.56	1.55	1.63	1.79	1.96	2.12	2.17	0.76
Interactive	[1.73]	[1.55]	[1.47]	[1.56]	[1.78]	[1.85]	[1.99]	[2.06]	(0.14)
Nonroutine	1.64	1.60	1.49	1.84	1.78	1.94	1.88	1.95	1.19
Manual	[1.63]	[1.63]	[1.44]	[1.85]	[1.79]	[1.97]	[1.82]	[2.01]	(0.22)
Routine	0.20	0.16	0.12	0.14	0.13	0.08	0.11	0.07	1.02
Cognitive	[0.20]	[0.16]	[0.11]	[0.13]	[0.13]	[0.08]	[0.11]	[0.07]	(0.03)
Routine	0.20	0.17	0.15	0.14	0.09	0.04	0.02	0.01	1.03
Manual	[0.20]	[0.18]	[0.18]	[0.18]	[0.11]	[0.05]	[0.03]	[0.01]	(0.01)

Notes: See the notes for Table 5. In Table 5, we include not only the words mentioned in footnote 14, but also similar words according to our continuous bag of words model. Here, instead, we apply only the mapping between task groups and words mentioned in footnote 14.

Table 12: Trends in Keyword Frequencies

	Frequency								Within
	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	Share
Nonroutine	3.19	3.50	3.33	3.63	3.93	3.86	4.63	5.48	0.93
Analytic	[3.15]	[3.40]	[3.32]	[3.58]	[3.78]	[3.65]	[4.38]	[5.29]	(0.07)
Nonroutine	4.94	4.56	4.57	5.18	5.84	5.97	6.56	7.03	1.01
Interactive	[4.86]	[4.42]	[4.36]	[4.89]	[5.71]	[5.73]	[6.27]	[6.97]	(0.07)
Nonroutine	2.50	2.36	2.08	2.54	2.59	2.63	2.59	2.73	1.69
Manual	[2.43]	[2.33]	[2.04]	[2.52]	[2.49]	[2.58]	[2.52]	[2.82]	(0.64)
Routine	1.00	0.91	0.69	0.65	0.74	0.71	0.59	0.54	0.98
Cognitive	[1.01]	[0.91]	[0.68]	[0.66]	[0.75]	[0.80]	[0.63]	[0.57]	(0.06)
Routine	0.70	0.58	0.45	0.42	0.27	0.15	0.10	0.07	0.99
Manual	[0.77]	[0.59]	[0.51]	[0.57]	[0.33]	[0.21]	[0.14]	[0.15]	(0.08)

Notes: See the notes for Table 5. In comparison we here apply an occupation classification scheme based on 6-digit SOC codes, as opposed to 4-digit SOC codes.

Table 13: Trends in Keyword Frequencies

	Frequency								Within
	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95-99	Share
Nonroutine	4.06	4.47	4.26	4.93	4.77	4.63	5.36	6.42	0.68
Analytic	[3.95]	[4.22]	[3.97]	[4.34]	[4.54]	[4.35]	[4.93[	[5.54]	
Nonroutine	5.50	4.84	5.07	5.67	6.31	6.60	7.42	8.10	0.59
Interactive	[5.30]	[4.85]	[4.83]	[5.11]	[5.89]	[6.10]	[6.40]	[6.84]	
Nonroutine	1.75	1.71	1.61	1.92	2.03	2.20	2.20	2.15	0.89
Manual	[1.77]	[1.71]	[1.60]	[1.82]	[1.99]	[2.11]	[2.20]	[2.13]	
Routine	1.47	1.33	1.03	0.96	0.99	0.85	0.55	0.61	0.60
Cognitive	[1.51]	[1.38]	[1.15]	[1.15]	[1.20]	[1.17]	[1.00]	[1.00]	
Routine	0.45	0.38	0.27	0.28	0.18	0.10	0.05	0.03	0.92
Manual	[0.46]	[0.37]	[0.29]	[0.29]	[0.20]	[0.13]	[0.10]	[0.08]	

Notes: See the notes for Table 5. In comparison, we here apply an occupation classification scheme based on job titles, as opposed to 4-digit SOC codes.

D.3 Plots of Task Measures within Occupations

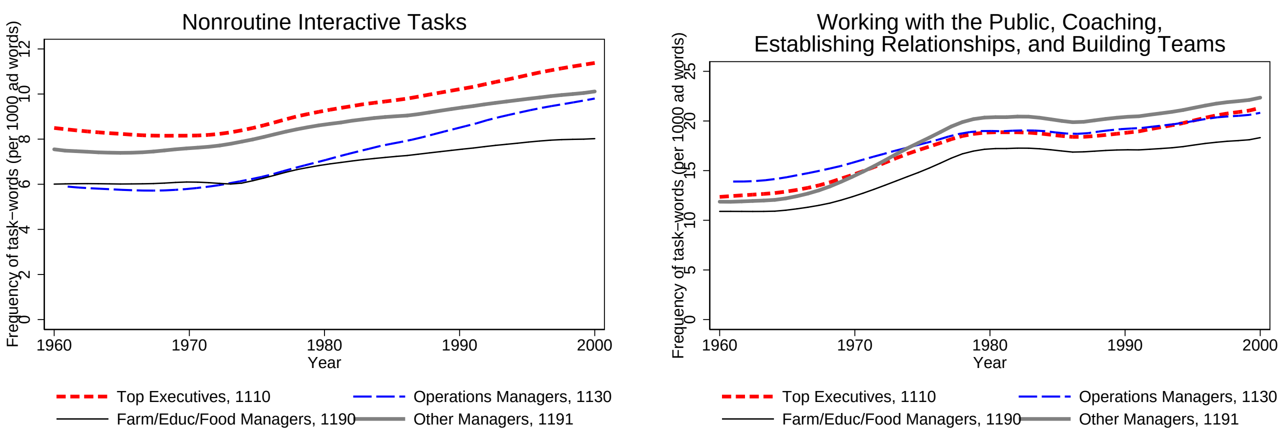
In this appendix, we plot changes in the frequencies of different task words within the three clusters of occupations that were the focus of Section 4.2. The goal of this appendix is to demonstrate that the trends which we had plotted in Section 4.2 reflect changes within 4-digit SOC occupations — as opposed to changes between 4-digit occupations — within these clusters of occupations. In Figure 20, we plot changes in the frequencies of nonroutine interactive tasks (left panel) and the interpersonal activities that were mentioned in the National Research Council (1999) report. Both sets of measures suggest that, within 4-digit occupations, tasks involving working and communicating with other people have become more important in managerial occupations.

Second, in Figure 21, we plot trends in routine cognitive tasks (left panel) and nonroutine interactive tasks (right panel) for three office clerk occupations. Within each of the three occupations, routine cognitive tasks have been mentioned in job ads less and less frequently over time; the largest decrease has occurred for financial clerks. On the other hand, mentions of nonroutine interactive tasks have displayed no trend for two occupations, but have increased considerably for the third (credit/file clerks).

Finally, Figure 22 presents the changes in routine manual tasks and nonroutine analytic tasks for six separate production worker occupations. Across all six occupations, the frequency of routine manual words has declined. With the exception of the "Printing" occupation, the frequency of words related to nonroutine analytic tasks in job ads for each of these occupations has increased.

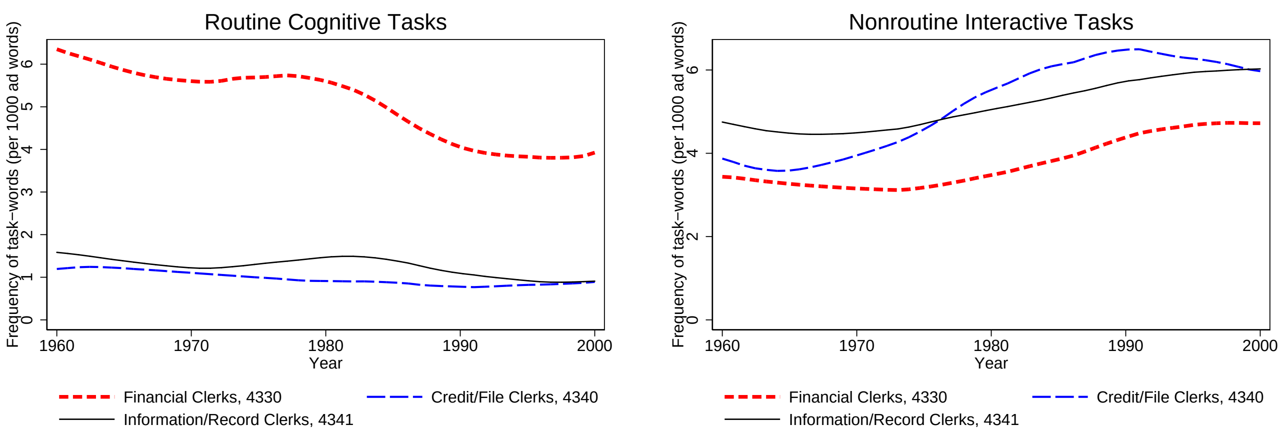
In summary, Figures 20 to 22 demonstrate that our narratives from Section 4.2 describe

Figure 20: Task Measures: Managerial Occupations



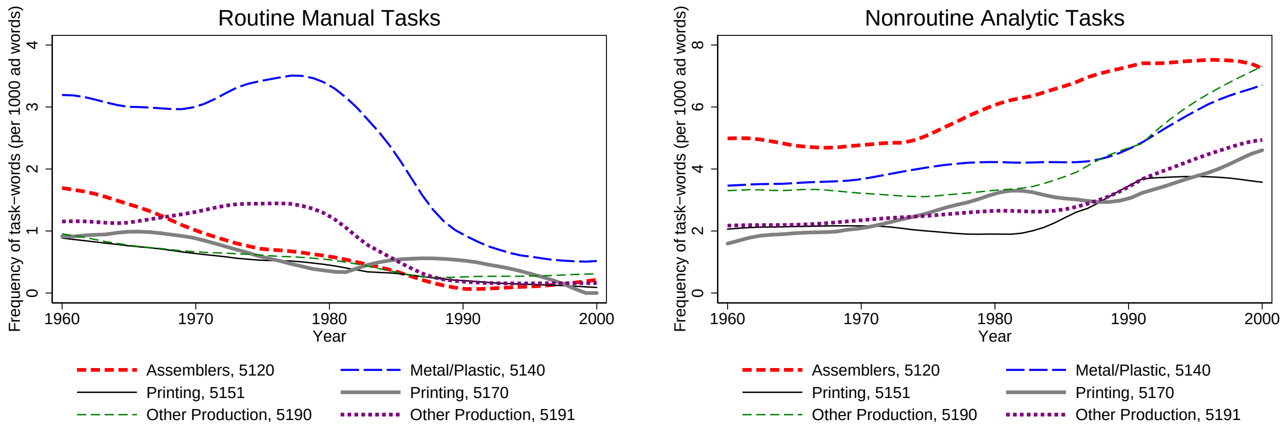
Notes: We apply a local polynomial smoother.

Figure 21: Task Measures: Office Clerk Occupations



Notes: We apply a local polynomial smoother.

Figure 22: Task Measures: Non-Supervisory Production Occupations



Notes: We apply a local polynomial smoother.

within-occupation changes in task content. Again, these changes would be difficult to detect using previous, conventional data sets tracking job content.

E Sensitivity Analysis and Supplemental Figures and Tables Related to Section 5.1

E.1 RIF Regression Results

In this appendix, we present supplemental tables and figures related to our wage distribution decompositions in Section 5.1. Tables 14 through 18 give results from the RIF regressions of male workers' log wage income in 1960, 1980, and 2000 (we omitted the results from 1970 and 1990 so as to fit each table on a single page). These RIF regressions correspond to the decompositions plotted in Figures 7 to 10 in Section 5.1. Each of the five tables use different measures of workers' task intensities. In Tables 14 and 15, we report the RIF regression coefficients using the task measures defined by Spitz-Oener (2006). In Table 14, the task measures freely vary throughout the sample period, while in Table 15 occupations' task measures are fixed at their sample means.

Tables 16 through 18 give the RIF coefficient estimates using Firpo, Fortin, and Lemieux (2014)'s task measures instead. The first two tables use text from the *Boston Globe*, *New York Times*, and *Wall Street Journal* to characterize occupational task intensities. In Table 16, the task measures are allowed to vary freely across time. In Table 17, we aim to isolate the differences in the final results that are due to the data source (O*NET versus newspaper data) as opposed to using occupational task measures which are changing over time. In other words, occupations' task intensities are fixed across the sample. Table 18 is closest, in its task measures, to Appendix Table A3 of Firpo, Fortin, and Lemieux (2014); it applies O*NET's task measures.

E.2 Decompositions with Alternate Measures, Samples

In this appendix, we begin to explore the sensitivity of our Section 5.1 results to our specification of tasks in our RIF regressions and to the sample. We begin in Figures 23 and 24 by modifying the way in which we measure our groups of nonroutine and routine tasks. In these two figures, we search only for the question titles in Spitz-Oener (2006)'s original definitions when constructing our task measures. Our measures do not include similar words identified from our continuous bag of words model.

Turning to the results, as the solid line of Figure 23 implies, decompositions in which the task measures are fixed within occupations account for a minor portion of the increase in inequality. This is consistent with the results given in Section 5.1, in our benchmark specification. Task measures which are allowed to vary throughout the sample period, within occupations, account for a 8 log point increase in 90-10 earnings inequality that has occurred between 1960 and 2000. This 8 log point figure is smaller than the 26 log point effect which we report in Section 5.1.

In Figure 24, we plot the composition and wage structure effects separately. The primary source of the difference between Figure 23 and Figure 7 is in the composition, as opposed to the wage structure, effects. According to the left panel of Figure 24, the composition effects account for a 29 log point increase in 90-10 inequality, similar to the 32 log point increase reported in the left panel of Figure 8. However, the (unobserved) wage structure effects account for a countervailing 20 log point decrease in 90-10 inequality, bigger than the 6 log point decrease reported in the right panel of Figure 8. As in Section 5.1, changes in task composition of the workforce, as opposed to changes in the return to tasks, are the primary source of the increase in earnings inequality.

Throughout Section 5.1, we transformed the raw newspaper text frequencies using the normalization: $T_{hjt} = \left(\widetilde{T}_{hjt} - \mu_h\right)/\sigma_h$, where, as a reminder, μ_h is the mean and σ_h is the standard deviation of keyword frequencies for task group h across years and occupations. As a robustness check, in Figure 25 we consider decompositions which are based on an alternate normalization. Instead of subtracting by the mean and dividing by the standard deviation, we instead rank newspaper ads according to the number of mentions of each task: $T_{hjt} = \mathrm{rank_h}\widetilde{T}_{hjt}$. This measure ranges between -0.5 (for ads that have the minimum mentions of the given task compared to all other newspaper ads in our newspaper data) to

Table 14: RIF Regression Coefficients on Male Log Earnings, Spitz-Oener (2006) Task Measures

Year	1960		1980		20000	
Quantile	10	90	10	90	10	90
Non white	-0.505	-0.123	-0.194	-0.088	-0.076	-0.093
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.002)
Not married	-0.652	-0.112	-0.657	-0.059	-0.512	-0.136
	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.002)
Experience<10	-0.586	-0.451	-0.607	-0.356	-0.948	-0.299
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)
10≤Experience<20	-0.035	-0.142	-0.051	-0.159	-0.045	-0.094
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Experience≥30	-0.012	-0.003	0.049	-0.029	-0.048	0.014
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Some High School	-0.230	-0.342	-0.512	-0.156	-0.775	-0.056
	(0.003)	(0.003)	(0.004)	(0.001)	(0.006)	(0.002)
High School	0.002	-0.198	0.030	-0.071	-0.017	-0.067
	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)	(0.001)
College	0.176	0.497	0.294	0.285	0.310	0.458
	(0.003)	(0.005)	(0.003)	(0.002)	(0.003)	(0.003)
Post-grad	0.149	0.614	0.252	0.434	0.266	1.017
	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)
Nonroutine Analytic	-0.010	0.083	0.067	0.045	0.122	0.038
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nonroutine Interactive	0.001	0.127	-0.011	0.088	0.007	0.124
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Routine Cognitive	0.095	-0.007	0.014	-0.020	0.097	0.023
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Routine Manual	0.049	-0.021	0.099	-0.004	0.091	-0.061
	(0.001)	(0.000)	(0.001)	(0.000)	(0.006)	(0.002)
Nonroutine Manual	-0.041	-0.035	-0.006	-0.005	0.058	-0.023
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.000)
Constant	9.668	11.192	9.985	11.317	10.139	11.380
	(0.002)	(0.003)	(0.003)	(0.002)	(0.004)	(0.002)
N	1423710	1423710	1841120	1841120	2223145	2223145
Adjusted R^2	0.175	0.171	0.128	0.146	0.133	0.167

Notes: The table contains coefficients and standard errors from RIF regressions. The task measures are computed using data from the *Boston Globe*, *New York Times*, and *Wall Street Journal*, with keyword frequencies computed on a year-by-year basis and then smoothed using a locally weighted polynomial regression. Each of the four activity measures are normalized to have mean 0 and standard deviation 1, averaged over all individuals in the regression sample.

Table 15: RIF Regression Coefficients on Male Log Earnings, Spitz-Oener (2006) Task Measures

Year	1960		1980		2000	
Quantile	10	90	10	90	10	90
Non white	-0.507	-0.119	-0.191	-0.087	-0.077	-0.093
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.002)
Not married	-0.651	-0.103	-0.658	-0.059	-0.504	-0.135
	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.002)
Experience<10	-0.589	-0.435	-0.609	-0.356	-0.947	-0.310
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)
10≤Experience<20	-0.036	-0.139	-0.051	-0.159	-0.047	-0.098
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Experience≥30	-0.009	-0.006	0.049	-0.029	-0.047	0.013
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Some High School	-0.239	-0.309	-0.507	-0.156	-0.803	-0.051
	(0.003)	(0.003)	(0.004)	(0.001)	(0.006)	(0.002)
High School	-0.002	-0.173	0.031	-0.071	-0.035	-0.059
	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)	(0.001)
College	0.173	0.448	0.298	0.287	0.333	0.441
	(0.003)	(0.005)	(0.003)	(0.002)	(0.003)	(0.003)
Post-grad	0.150	0.555	0.262	0.439	0.297	0.998
	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)
Nonroutine Analytic	-0.003	0.131	0.048	0.037	0.119	0.048
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Nonroutine Interactive	0.015	0.117	0.011	0.084	0.007	0.154
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Routine Cognitive	0.089	-0.010	0.044	-0.018	0.044	-0.025
	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.001)
Routine Manual	0.069	-0.018	0.077	-0.002	0.098	-0.010
	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)
Nonroutine Manual	-0.026	-0.027	0.009	-0.008	0.057	-0.026
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Constant	9.735	11.109	9.974	11.299	10.122	11.463
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
N	1423710	1423710	1841120	1841120	2223145	2223145
Adjusted R^2	0.174	0.180	0.127	0.146	0.136	0.170

Notes: The table contains coefficients and standard errors from RIF regressions. The task measures are computed using data from the *Boston Globe*, *New York Times*, and *Wall Street Journal*, with keyword frequencies computed as averages over the 1960-2000 period. Each of the four activity measures are normalized to have mean 0 and standard deviation 1, averaged over all individuals in the regression sample.

Table 16: RIF Regression Coefficients on Male Log Earnings, Firpo, Fortin, and Lemieux (2014) Task Measures

Year	1960		1980		2000	
Quantile	10	90	10	90	10	90
Non white	-0.503	-0.135	-0.194	-0.095	-0.082	-0.110
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.002)
Not married	-0.628	-0.090	-0.659	-0.058	-0.513	-0.138
	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.002)
Experience<10	-0.555	-0.417	-0.604	-0.354	-0.956	-0.303
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)
10≤Experience<20	-0.030	-0.136	-0.048	-0.159	-0.048	-0.096
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Experience≥30	-0.015	-0.008	0.048	-0.028	-0.046	0.020
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Some High School	-0.192	-0.348	-0.472	-0.177	-0.779	-0.102
	(0.003)	(0.003)	(0.004)	(0.001)	(0.006)	(0.002)
High School	0.030	-0.192	0.052	-0.082	-0.016	-0.093
	(0.003)	(0.003)	(0.003)	(0.001)	(0.003)	(0.001)
College	0.173	0.448	0.280	0.292	0.316	0.493
	(0.003)	(0.005)	(0.003)	(0.002)	(0.003)	(0.003)
Post-grad	0.178	0.561	0.238	0.410	0.281	1.017
	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)
Information Content	-0.283	-0.117	0.044	-0.063	0.066	-0.026
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Face-to-Face	0.263	0.053	0.147	-0.041	-0.015	0.029
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)
On-Site Job	-0.009	-0.081	0.030	0.002	0.137	-0.053
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Decision-Making	-0.078	0.328	-0.116	0.175	0.042	0.084
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)
Constant	9.608	11.270	9.955	11.329	10.121	11.398
	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)
N	1423710	1423710	1841120	1841120	2223145	2223145
Adjusted R^2	0.184	0.179	0.126	0.143	0.133	0.158

Notes: The table contains coefficients and standard errors from RIF regressions. The task measures are computed using data from the *Boston Globe*, *New York Times*, and *Wall Street Journal*, with keyword frequencies computed on a year-by-year basis and then smoothed using a locally weighted polynomial regression. Each of the four activity measures are normalized to have mean 0 and standard deviation 1, averaged over all individuals in the regression sample.

Table 17: RIF Regression Coefficients on Male Earnings, Firpo, Fortin, and Lemieux (2014) Task Measures

Year	1960		1980		2000	
Quantile	10	90	10	90	10	90
Non white	-0.444	-0.138	-0.187	-0.095	-0.076	-0.109
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.002)
Not married	-0.602	-0.092	-0.644	-0.058	-0.499	-0.135
	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.002)
Experience<10	-0.516	-0.412	-0.590	-0.353	-0.945	-0.303
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)
10≤Experience<20	-0.026	-0.133	-0.047	-0.158	-0.048	-0.098
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Experience≥30	-0.016	-0.007	0.046	-0.028	-0.047	0.017
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Some High School	-0.192	-0.342	-0.470	-0.175	-0.764	-0.082
	(0.003)	(0.003)	(0.004)	(0.001)	(0.006)	(0.002)
High School	0.015	-0.190	0.048	-0.080	-0.010	-0.075
	(0.002)	(0.003)	(0.003)	(0.001)	(0.003)	(0.001)
College	0.152	0.432	0.281	0.289	0.321	0.448
	(0.003)	(0.005)	(0.003)	(0.002)	(0.003)	(0.003)
Post-grad	0.102	0.482	0.242	0.413	0.291	0.952
	(0.004)	(0.007)	(0.003)	(0.003)	(0.004)	(0.005)
Information Content	-0.056	-0.071	-0.006	-0.056	-0.003	-0.042
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Face-to-Face	0.196	-0.014	0.127	-0.036	0.168	-0.026
	(0.002)	(0.002)	(0.003)	(0.001)	(0.003)	(0.002)
On-Site Job	-0.044	-0.022	0.035	-0.006	0.113	-0.055
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Decision-Making	-0.141	0.252	-0.083	0.155	-0.087	0.195
	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.002)
Constant	9.743	11.130	9.954	11.301	10.107	11.473
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
N	1427801	1427801	1846332	1846332	2230601	2230601
Adjusted R^2	0.161	0.180	0.126	0.146	0.136	0.162

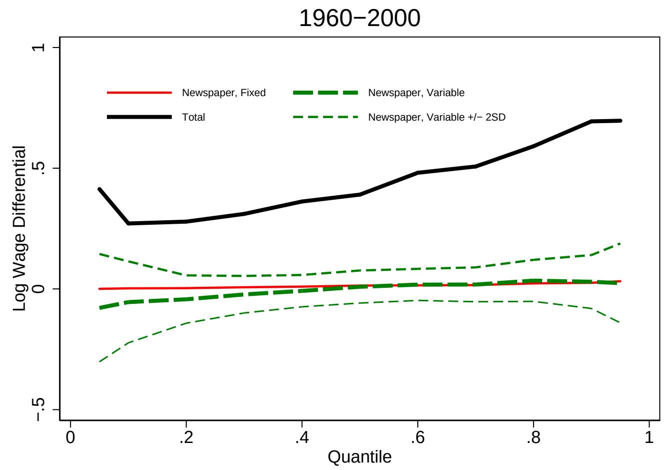
Notes: The table contains coefficients and standard errors from RIF regressions. The task measures are computed using data from the *Boston Globe*, *New York Times*, and *Wall Street Journal* with keyword frequencies computed as averages over the 1960-2000 period. Each of the four activity measures are normalized to have mean 0 and standard deviation 1, averaged over all individuals in the regression sample.

Table 18: RIF Regression Coefficients on Male Log Earnings, Firpo, Fortin, and Lemieux (2014) Task Measures

Year	1960		1980		2000	
Quantile	10	90	10	90	10	90
Non white	-0.439	-0.105	-0.183	-0.087	-0.059	-0.096
	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)	(0.002)
Not married	-0.602	-0.087	-0.623	-0.053	-0.466	-0.130
	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.002)
Experience<10	-0.515	-0.405	-0.576	-0.348	-0.907	-0.293
	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.002)
10≤Experience<20	-0.023	-0.129	-0.043	-0.156	-0.043	-0.095
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Experience≥30	-0.017	-0.015	0.050	-0.031	-0.035	0.017
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Some High School	-0.186	-0.295	-0.469	-0.155	-0.769	-0.051
	(0.003)	(0.003)	(0.004)	(0.001)	(0.006)	(0.002)
High School	0.016	-0.170	0.039	-0.073	-0.021	-0.057
	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)	(0.001)
College	0.132	0.427	0.260	0.272	0.280	0.405
	(0.003)	(0.006)	(0.003)	(0.002)	(0.003)	(0.003)
Post-grad	0.102	0.455	0.207	0.396	0.184	0.906
	(0.004)	(0.007)	(0.003)	(0.003)	(0.004)	(0.005)
Information Content	-0.048	0.038	0.071	-0.010	0.134	-0.043
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Face-to-Face	0.073	-0.010	0.025	0.014	-0.048	0.038
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Automation/Routine	-0.025	-0.105	-0.101	-0.049	-0.177	-0.069
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
On-Site Job	-0.050	-0.109	0.052	-0.061	0.124	-0.148
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Decision-Making	0.076	0.209	0.122	0.116	0.220	0.180
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
Constant	9.721	11.119	9.921	11.296	10.053	11.469
	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)
N	1427801	1427801	1846332	1846332	2230601	2230601
Adjusted R^2	0.162	0.178	0.133	0.146	0.148	0.169

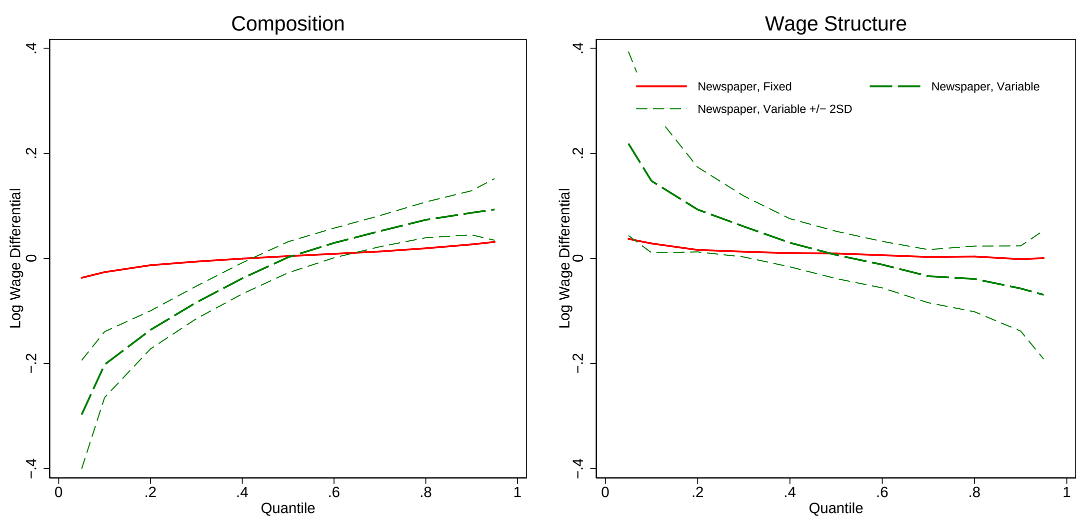
Notes: The table contains coefficients and standard errors from RIF regressions. The task measures are computed using data from O*NET and the definitions of Firpo, Fortin, and Lemieux (2014). Each of the five task measures are normalized to have mean 0 and standard deviation 1, averaged over all individuals in the regression sample.

Figure 23: Decomposition of Real Log Earnings: Specification of Task Measures



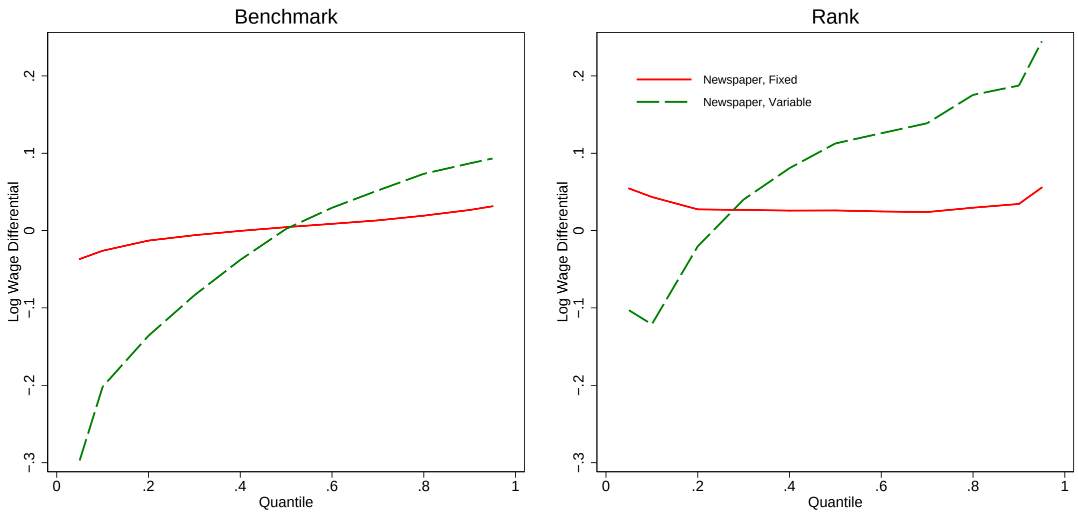
Notes: The figure describes the contribution of occupational characteristics, through both compositional and wage structure changes using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks. In Figure 7, we include not only the words mentioned in footnote 14, but also similar words, according to our continuous bag of words model. Here, instead, we apply only the mapping between task groups.

Figure 24: Decomposition of Real Log Earnings: Specification of Task Measures



Notes: The left panel describes the contribution of occupational characteristics through compositional changes using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks. The right panel describes the contribution of occupational characteristics through wage structure effects. In Figure 7, we include not only the words mentioned in footnote 14, but also similar words, according to our continuous bag of words model. Here, instead, we apply only the mapping between task groups and question titles used in Spitz-Oener (2006).

Figure 25: Decomposition of Real Log Earnings: Specification of Task Measures



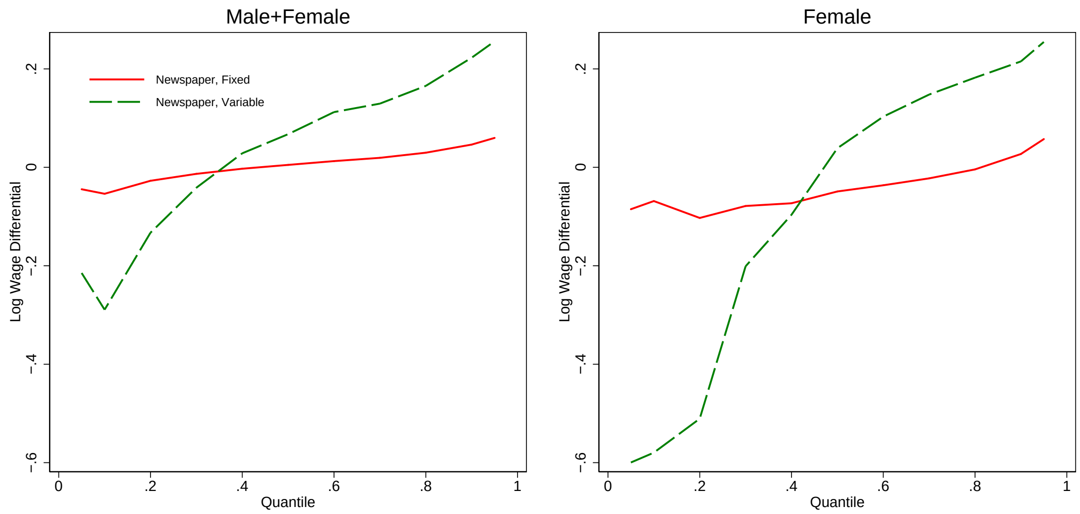
Notes: The figure describes the contribution of occupational characteristics, through both compositional and wage structure changes, using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks.

0.5 (for ads that mention the particular task most frequently). With this alternate definition, the contribution of our task measures to 90-10 inequality is somewhat larger — 31 log points versus 26 log points — than in our benchmark specification.

Up to now, our RIF regression-based decompositions included only male workers in our samples. We excluded female workers from our sample because of the vast increase in female labor force participation that has occurred since 1960, especially in the 1960s and 1970s. In Figure 26, we recompute our decompositions with both male and female workers (in the left panel) and female workers only (in the right panel). Qualitatively, the results are similar to those given in Figure 7. However, including female workers increases the measured role of our task measures in accounting for the increase in earnings inequality. According to the decomposition plotted in the left panel, Spitz-Oener (2006)'s task measures account for a 51 log point increase in earnings inequality. The difference between these decompositions and those using our benchmark sample are primarily due to the wage structure effect (unplotted) and are primarily seen in the decompositions spanning the earlier part of the sample.

In sum, our benchmark results, given in Figure 7 in Section 5.1, assigned a 26 log point contribution of our task measure to the increase in 90-10 inequality, primarily through composition as opposed to wage structure effects. In this appendix, we considered three robustness checks: to the way in which our task measures are measured, to the transformation applied to the raw newspaper frequency counts, and to the sample. In the first robustness

Figure 26: Decomposition of Real Log Earnings: Male and Female



Notes: The figure describes the contribution of occupational characteristics, through both compositional and wage structure changes, using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks.

check, the contribution of our task measures to earnings inequality is somewhat smaller than in our benchmark specification. In the latter two robustness checks, the contribution of our task measures is at least as large as those reported in Section 5.1. Moreover, as in Section 5.1, in each of these three robustness checks the primary contribution of our task measures is through composition effects.

E.3 Decompositions That Use Census Data from Only New York and Boston

Our newspaper data contain information almost exclusively about vacancies in the New York City and Boston metro areas. We used this information about New York City and Boston ads to characterize the skill and task content of jobs throughout the United States. This discrepancy could potentially be problematic, especially since workers in New York City and Boston are not representative of U.S. workers more generally. What is more, this non-representativeness may be growing over time (because, for example, the college graduate share in New York City and Boston has increased faster than in other parts of the country.) In this appendix, we re-compute some of our main figures, using only census data on workers from the New York City and Boston metropolitan statistical areas.[42] As a preview, the

[42]To ensure that the boundaries of the MSAs are fixed through time, we remove individuals who reside in counties which were added to the New York MSA definitions part of the way through our sample period:

main results of this appendix are that i) the relationship between occupations' shares of newspaper vacancy posting counts and their shares of workers in the decennial census is slightly stronger when we restrict our census data to only include individuals residing in the New York and Boston metro areas and that ii) the RIF-based wage decompositions paint a similar picture for the role of technology in accounting for labor income inequality. While we prefer the paper's main sample to include workers throughout the United States, since this choice allows us to make statements about the U.S. wage distribution (an object of greater intrinsic interest than the wage distribution of workers residing in New York City and Boston), because of the discrepancy between the regions covered by the newspaper text and decennial census samples we are admittedly ambivalent about this choice.

Figures 27 and 28 present the share of workers in each occupation in the decennial census and the share of vacancy postings in each occupation in our newspaper data. Again, compared to Figures 13 and 14, the sole difference is that we are looking at workers who in the census are reported as living in either the Boston MSA or the New York MSA. The two data sources now align slightly more closely here, than in Figures 13 and 14. Across 2-digit occupations, the correlation of the shares depicted in Figure 27 are 0.77, 0.92, and 0.72 for 1960, 1980, and 2000, respectively — on average 13 percentage points higher than in comparable years when using census data from the entire U.S. In Figure 28, the correlations among occupations' newspaper vacancy postings and occupations New York and Boston worker shares at the 4-digit level are 0.48, 0.61, 0.59, and 0.50 for 1960, 1970, 1980, and 2000, respectively.

Next, we examine whether the wage decompositions differ when using only census data from the New York and Boston metro areas. Within the New York and Boston metro areas, overall labor earnings inequality increased more rapidly than in the U.S. overall. Between 1960 and 2000, the 90-10 difference of log earnings increased by 75 log points; the 50-10 difference of log earnings increased by 43 log points. While the increase in inequality is greater in Boston and New York City than in the country as a whole, as in Section 5.1, Spitz-Oener (2006)'s measures explain a 24 log point increase in 90-10 earning inequality, a 21 log point increase in the 50-10 difference.

In Figure 30, we plot the composition and wage structure effects separately. The task measures, through changes in composition, account for a 19 log point portion of the increase in 90-10 inequality, with a 2-standard deviation confidence interval of 14 and 25 log points for this effect.

The detailed decomposition (Figure 31) indicates that changes in the composition of

Hunterdon County, Middlesex County, Somerset County, Sussex County, and Warren County. All five of these counties are in New Jersey.

Figure 27: Occupation Shares

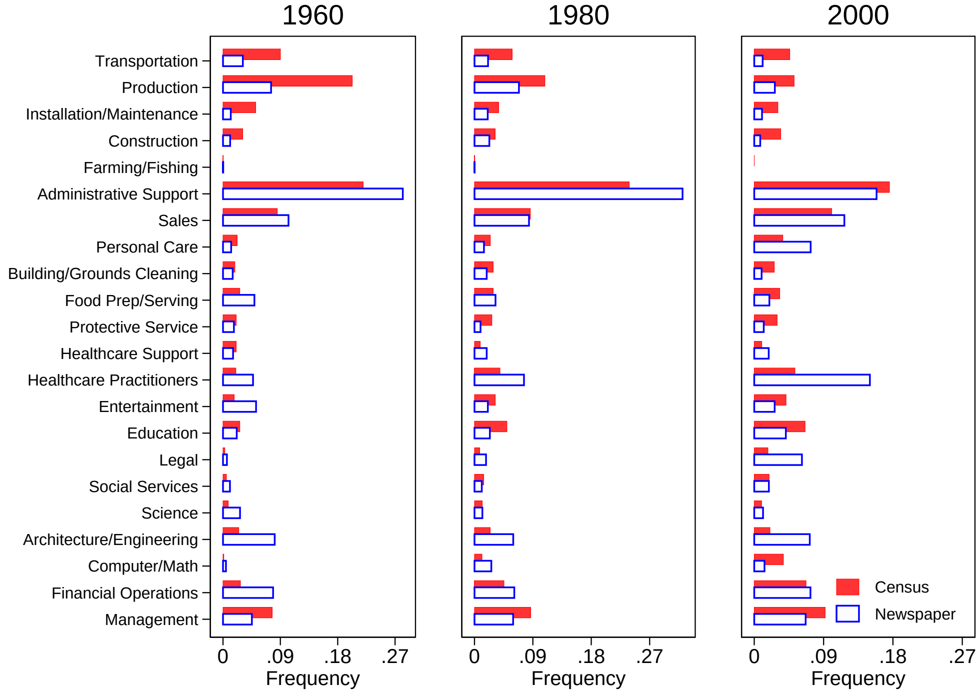


Figure 28: Occupation Shares

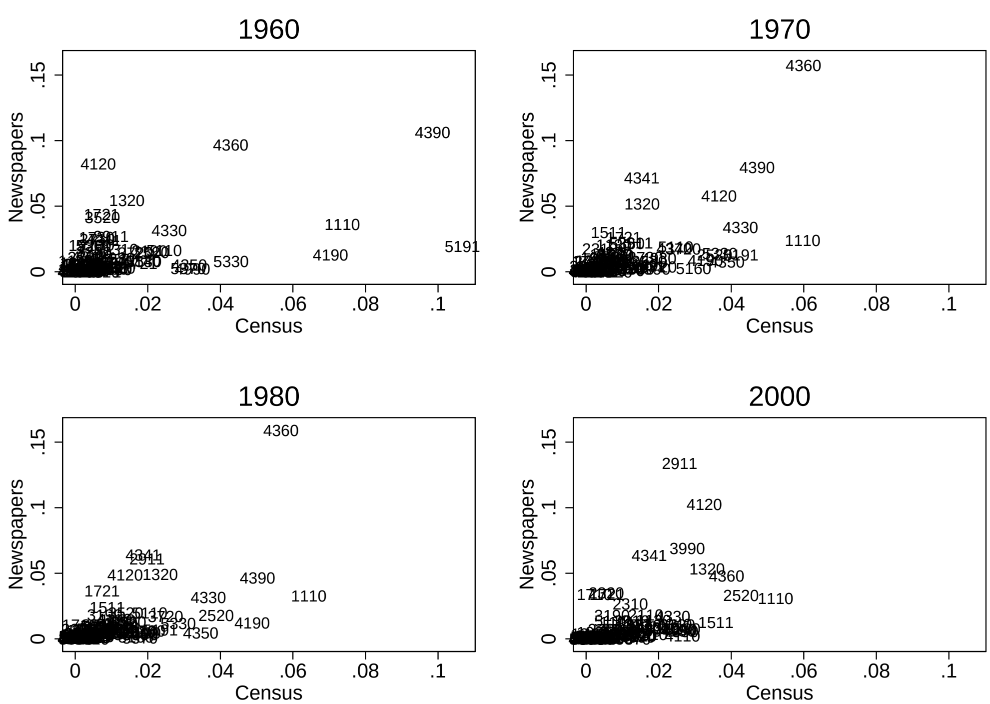
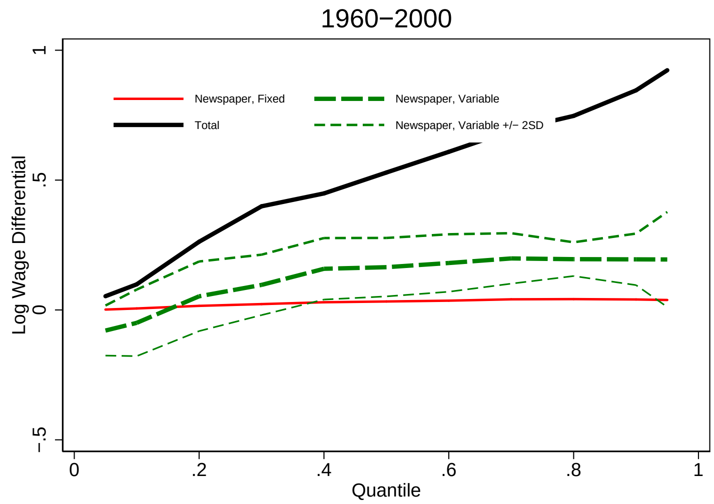
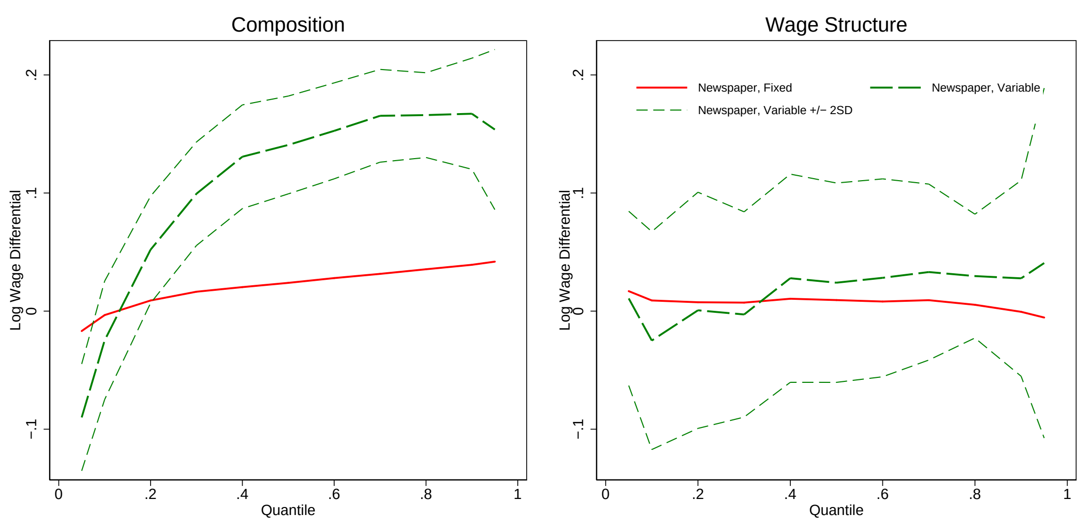


Figure 29: Detailed Decompositions: Boston and New York MSA Data



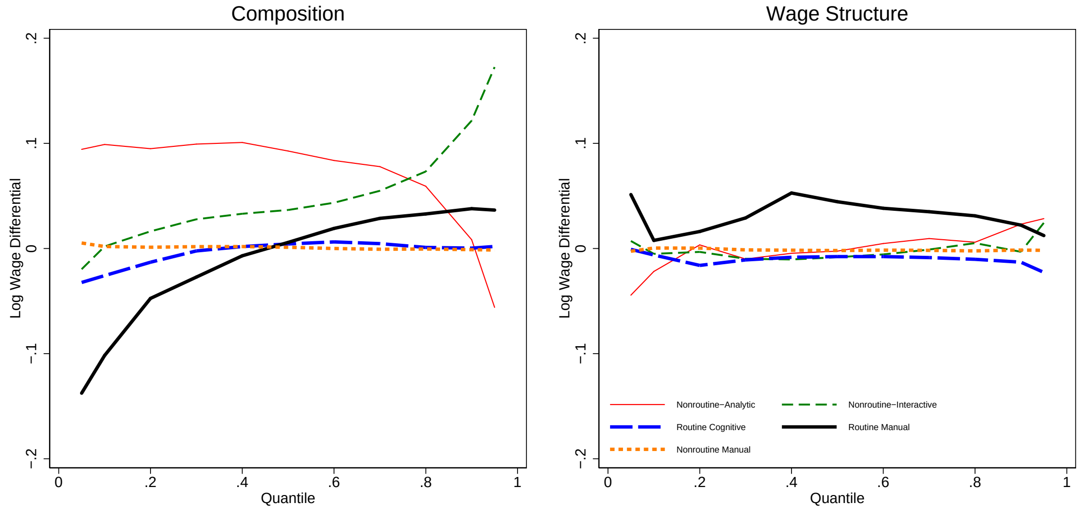
Notes: The figure describes the contribution of occupational characteristics, through both compositional and wage structure changes, using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks.

Figure 30: Decomposition of Real Log Earnings: Composition and Wage Structure, Boston and New York MSA Data



Notes: The left panel describes the contribution of occupational characteristics through compositional changes using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks. The right panel describes the contribution of occupational characteristics through wage structure effects. The individual components of the composition and wage structure effects can be found in Figure 31.

Figure 31: Detailed Decompositions: Boston and New York MSA Data



Notes: The panels describe the contribution of individual occupational characteristics, through compositional changes (left panel) and wage structure effects (right panel), to changes in the wage distribution. In this figure, we use the newspaper-based task measures that are allowed to vary within occupations across time.

jobs, as measured by their nonroutine interactive and routine manual task measures account for a 11 log point and 14 log point increase in 90-10 inequality. The contribution of nonroutine analytic tasks, through wage structure effects, explain a 4 log point increase in 90-10 inequality, with most of this effect occurring in the bottom half of the earnings distribution.

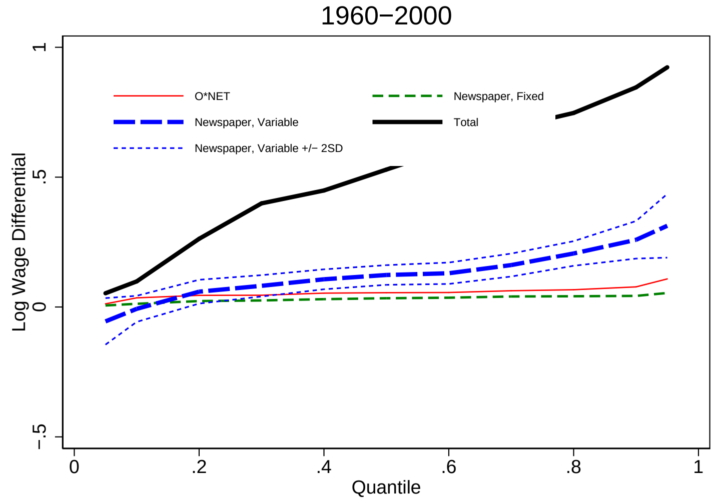
In Figure 32, we present the change in the distribution of earnings between 1960 and 2000, and the changes which are due to technological progress and offshoring, according to the Firpo, Fortin, and Lemieux (2014) categorization of O*NET Elements. The task measures, here, contribute a 27 log point increase in 90-10 inequality; the 2 standard deviation confidence interval spans 0.19 to 0.35.

Finally, in Figure 33, we break down the contribution of the composition and wage structure effects. Unlike in Section 5.1, the wage structure effects and composition effects account for a similar increase in labor earnings inequality.

E.4 Decompositions That Use the CPS Merged Outgoing Rotation Group

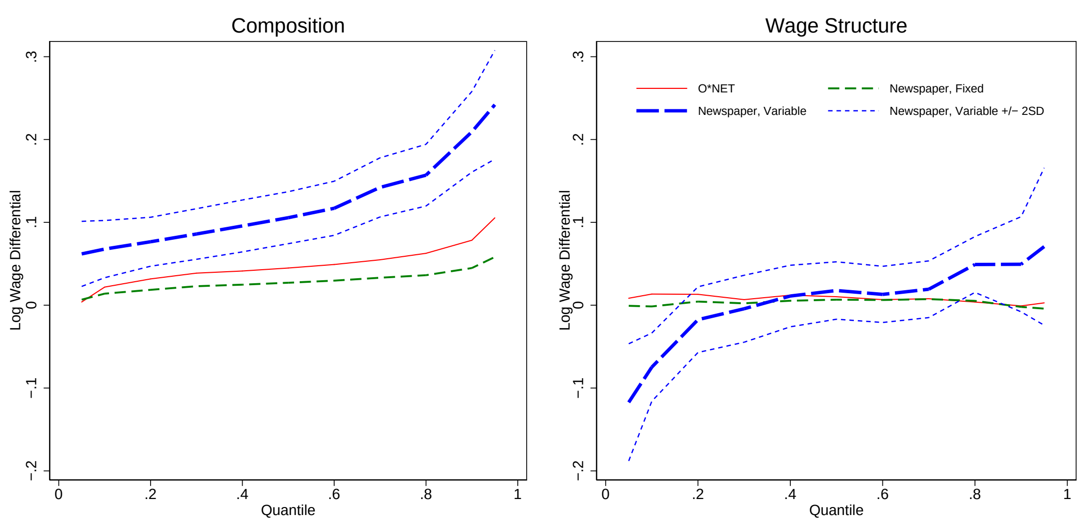
Firpo, Fortin, and Lemieux (2014) used the CPS Merged Outgoing Rotation Group (MORG) to form the sample for their RIF-based decompositions (these authors used the 1976-1978 May CPS in the first few years of the sample, to extend the period over which union status

Figure 32: Decomposition of Real Log Earnings: Boston and New York MSA Data



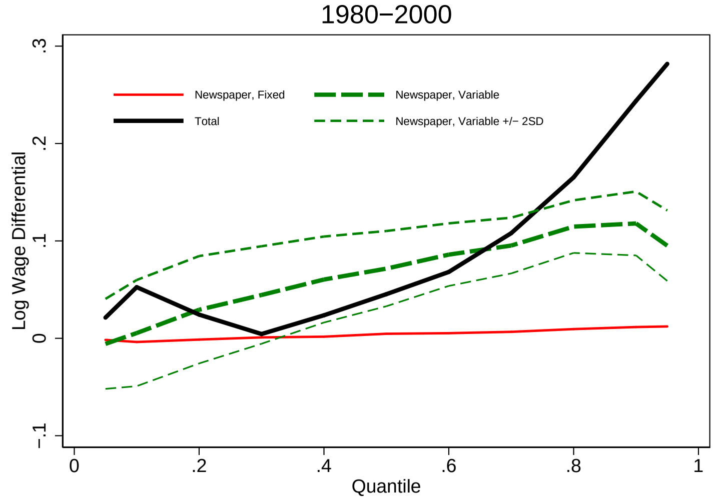
Notes: The thick solid line presents changes in log wage income of workers at different quantiles of the income distribution. The thin solid line and dashed lines give the contribution of occupations (both through the composition effects and wage structure effects) using three different measures of occupational characteristics.

Figure 33: Decomposition of Real Log Earnings: Composition and Wage Structure, Boston and New York MSA Data



Notes: The left panel describes the contribution of occupational characteristics through compositional changes, using Firpo, Fortin, and Lemieux (2014)'s O*NET based definitions. The right panel describes the contribution of occupational characteristics through wage structure effects. The individual components of the composition and wage structure effects can be found in Figure 29.

Figure 34: Decomposition of Real Log Hourly Wage: MORG Data



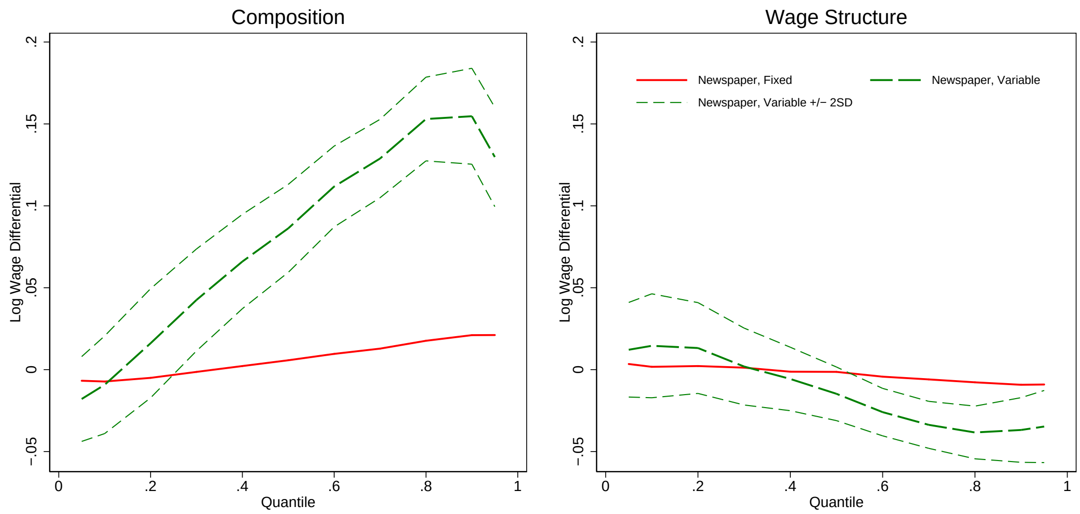
Notes: The figure describes the contribution of occupational characteristics, through both compositional and wage structure changes, using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks.

could be measured.) In our baseline analysis, we used the decennial census, leading us to use workers' labor income earnings as an outcome measure (as opposed to hourly wages). We chose the decennial census to extend the sample period as far back as our newspaper data would allow, while recognizing that total labor income earnings are not the ideal measure of the demand for individual workers' labor. In this appendix, we recompute Figures 7 to 10 using data from the 1980-2000 MORG. As in the body of the paper, but unlike in Appendix E.3, our sample of workers includes those from the entire U.S. Over this period, the 90-10 and 90-50 ratios of hourly workers increased by 19 log points and 22 log points respectively.

Figures 34, 35, and 36 present decomposition results, now using Spitz-Oener (2006)'s task measures. In Figure 34, we plot the change, via both composition and wage structure effects, accounted by our task measures. Our task measures account for a 11 log point increase in 90-10 wage inequality. This 11 log point contribution is equal to the total increase in earnings inequality, between 1980 and 2000, explained by our task measures when using our benchmark decennial census sample.

Figure 35 plots the composition and wage structure effects separately. The task measures, through composition effects, account for a 16 log point increase in 90-10 wage inequality (with a 2-standard deviation confidence interval of 12 to 21 log points). The wage effects account for a negligible portion of wage inequality. Figure 36 plots the composition and wage structure effects, separately for each of Spitz-Oener (2006)'s task measures. In Figure 36, we plot the results from the corresponding detailed decompositions. As in the benchmark specification,

Figure 35: Decomposition of Real Log Hourly Wage: Composition and Wage Structure, MORG Data



Notes: The left panel describes the contribution of occupational characteristics through compositional changes using Spitz-Oener (2006)'s definitions of routine and nonroutine tasks. The right panel describes the contribution of occupational characteristics through wage structure effects. The individual components of the composition and wage structure effects can be found in Figure 36.

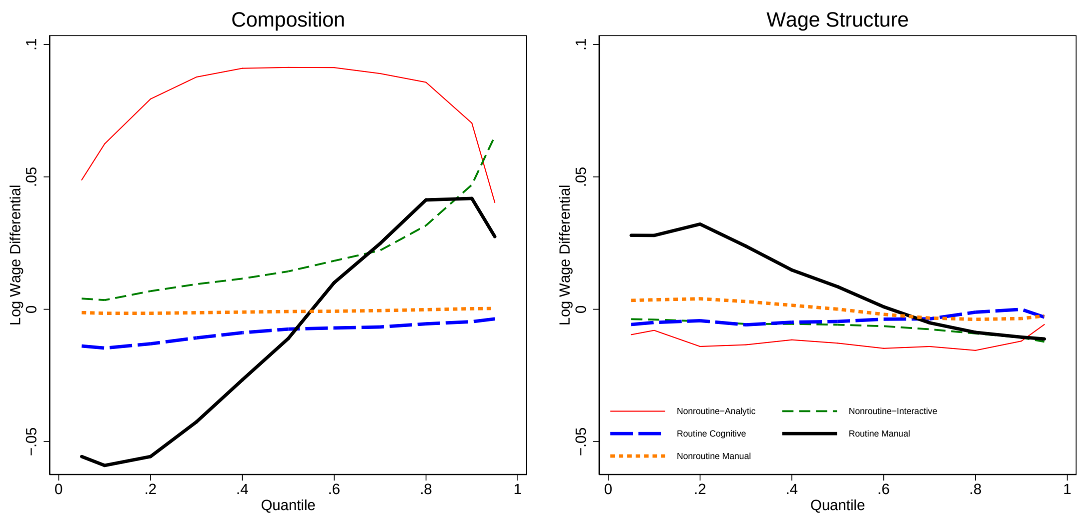
changes in the composition of workers' jobs (measured by their nonroutine interactive and routine manual tasks) are the primary contributors of increasing wage inequality.

In the remaining figures in this subsection, we re-compute our wage decompositions, now using Firpo, Fortin, and Lemieux (2014)'s occupational measures. Firpo, Fortin, and Lemieux (2014)'s measures of technological changes and offshorability account (combining both the composition and wage structure effects) for a 6 log point increase in the 90-10 ratio (with the two standard deviation confidence interval of 0.02 to 0.11) and a 5 log point increase in the 90-50 ratio of hourly wages, with the entire effects being due to the composition effects (see Figure 38, where we present the composition and wage structure effects separately). Decompositions based on time-invariant occupational measures account for at most a 1 percentage point increase in 90-10 or 90-50 inequality.

F Details of the Section 5.2 Model

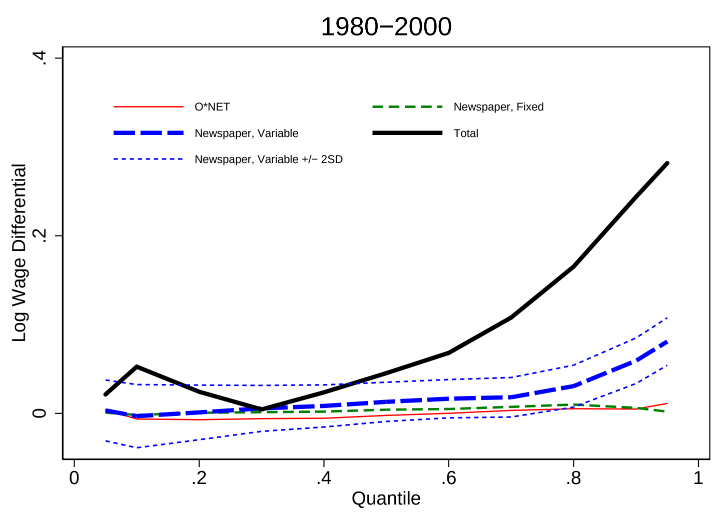
In this section, we describe the assumptions necessary to derive the empirical specification for our RIF regression decomposition from Equation 4. We then delineate two algorithms: In the the first, we compute the counterfactual changes in groups' wages, and in the second we apply those changes in groups' wages to compute changes in the overall distribution

Figure 36: Detailed Decompositions: MORG Data



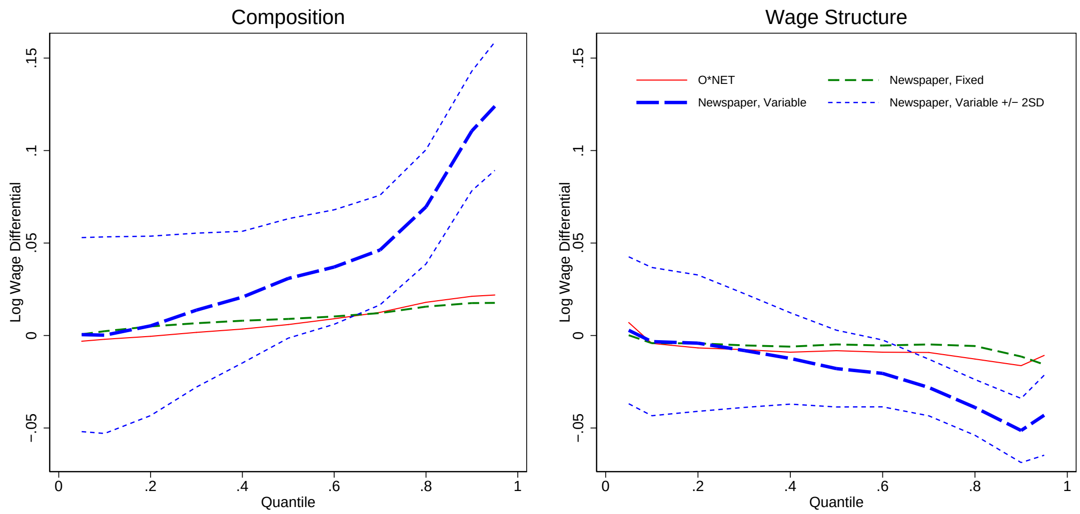
Notes: The panels describe the contribution of individual occupational characteristics, through compositional changes (left panel) and wage structure effects (right panel) to changes in the wage distribution. In this figure, we use the newspaper-based task measures which are allowed to vary within occupations across time.

Figure 37: Decomposition of Real Log Hourly Wage: MORG Data



Notes: The thick solid line presents changes in log wage income of workers at different quantiles of the income distribution. The thin solid line and dashed lines give the contribution of occupations (both through the composition effects and wage structure effects) using three different measures of occupational characteristics.

Figure 38: Decomposition of Real Log Hourly Wage: Composition and Wage Structure, MORG Data



Notes: The left panel describes the contribution of occupational characteristics through compositional changes using Firpo, Fortin, and Lemieux (2014)'s O*NET based definitions. The right panel describes the contribution of occupational characteristics through wage structure effects.

(within-group and between group) of wages. Throughout this section use $\hat{X}$ to refer to the ratio of variable X from one period to the next: $\hat{X} = \frac{X_{t+\tau}}{X_t}$.

Linking the Model of Occupations as a Bundle of Tasks to the RIF Regressions

In motivating our RIF regression based decomposition, we begin by relating the price of occupation-j specific output with the prices of its constituent tasks:[43]

$$\log P_{jt} = \log \pi_{0jt} + \sum_{h=1}^{H} T_{hjt} \log \pi_{hjt}. \tag{11}$$

In our model, perfect competition and homogeneous occupational output ensure that there is a single price for each occupation. Thus, although task prices π_{hjt} may vary by group reflecting the marginal value of each task in a production unit, these prices do not contribute to inequality once one takes into account the output price P_j. In other words, task prices only generate inequality between occupations. In contrast, the RIF regression specification, which allows for different task prices at different quantiles of the income distribution. One way to

[43]This is akin to the formulation in Yamaguchi (2012). To motivate this equation, suppose that workers cannot "un-bundle" their tasks in the sense of Heckman and Scheinkman (1987). To the extent that tasks cannot be "unbundled," task prices will differ across occupations.

justify this, more flexible specification, is to assume that workers at different quantiles work in different occupations. Another is to consider a more flexible model where the unobserved shocks are not Hicks neutral in a given occupation, but rather task-specific.

Bearing in mind that additional flexibility, we can use the pricing Equation 11 in the wage Equation 4 to obtain an expression for wages as a function of tasks, task prices, and skills:

$$\log W_{ijt} = \log \pi_{0jt} + \sum_{h=1}^{H} T_{hjt} \log \pi_{hjt} + \sum_{h=1}^{H} T_{hjt} \log S_{gh} + \log \epsilon_{ijt}. \tag{12}$$

We next assume that individual i's ability to perform task h at time t is linearly related to a set K of observable characteristics so that the underlying skill of group g individuals in performing task h can be written as $\log S_{gh} = \sum_{k=1}^{K} b_{hk} \tilde{S}_{gk}$. With this assumption, Equation 12 implies:

$$\log W_{ijt} = \log \pi_{0jt} + \sum_{h=1}^{H} T_{hjt} \log \pi_{hjt} + \sum_{k=1}^{K} \alpha_{kjt} \tilde{S}_{kg} + \log \epsilon_{ijt}, \tag{13}$$

where, again, $\tilde{S}_{gk}$ is an observable skill characteristic k for an individual in group g.

Equation 2, in the body of the paper, would be equivalent to Equation 13 if one were to remove "j" subscripts on the $\pi$and α terms. But, insofar as workers in different occupations are concentrated at different points in the wage distribution, our coefficient estimates from Equation 2 will vary at different quantiles.

Algorithm to Compute Counterfactual Changes in Wages

1. Start with an occupational price guess $\hat{P}^{[0]}$.

2. Compute

$$\hat{W}_{gj}^{[1]} = \hat{P}_j^{[0]} \prod_h (S_{hi})^{T_{hj}\left(\hat{T}_{hj}-1\right)}$$

$$\hat{\lambda}_{gj}^{[1]} = \frac{\left(\hat{W}_{gj}^{[1]}\right)^{\theta}}{\sum_{j'} \lambda_{gj'} \left(\hat{W}_{gj}^{[1]}\right)^{\theta}}$$

$$\widehat{\bar{W}_g}^{[1]} = \left(\sum_j \left(\hat{W}_{gj}^{[1]}\right)^{\theta} \lambda_{gj}\right)^{1/\theta}$$

3. Compute the excess demand function

$$
\begin{aligned}
Z_j^{[1]} &= \hat{E}^{[1]} - \sum_{g=1}^{G} \widehat{\bar{W}_g}^{[1]} \hat{\lambda}_{gj}^{[1]} \chi_{gj} \\
&= \sum_{g=1}^{G} \widehat{\bar{W}_g}^{[1]} \xi_g - \sum_{g=1}^{G} \widehat{\bar{W}_g}^{[1]} \hat{\lambda}_{gj}^{[1]} \chi_{gj}
\end{aligned}
$$

4. Compute the new prices

$$\hat{P}_j^{[1]} = \hat{P}_j^{[0]} + \nu Z_j^{[1]}$$

for some small adjustment number ν. We also need to normalize $P_1^{[1]} = 1$.

5. Check for convergence. If there is no convergence, go back to 2.

Algorithm to Simulate the Distribution of Wages

Given $\bar{W}_{g,1960}$ and $\widehat{\bar{W}_g}$ (constructed from the previous algorithm), we now know $\bar{W}_{g,t}$. Note that the distribution of wages for each type g in this model is Frechet, with parameters $\left(\bar{W}_{g,t}, \theta\right)$. To get at the distribution of wages within the model:

1. Fix a large number $\bar{L}^{sim}$ that controls the size of the simulation. We use $\bar{L}^{sim} = 10^6$.

2. For each group g, sample $\bar{L}^{sim} \times \sum_{j=1}^{J} \lambda_{gj}$ wages. To do so, draw for each observation in group i a unit exponential $\mathcal{U}$ and then apply the transformation $W = \bar{W}_{g,t} \cdot \mathcal{U}^{-1/\theta}$.

3. The vector of all W is a sample of wages in this economy.