Online Appendix

B Additional Calculations Related to Section I

This appendix presents comparisons of measures from our new data set to measures from the decennial census (Appendix B.1), from O*NET (Appendix B.2), and from the DOT (Appendix B.3). Then, we compute the top occupations for each Deming and Kahn (2018) skill measure (Appendix B.4).

B.1 Comparison to the Decennial Census

Certain variables are present in both our new newspaper data and in the census. 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.

First, Figures 9 and 10 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 9 presents this relationship at the 2-digit level. In the six decades depicted within the figure, the correlations between census job frequencies and frequencies in our newspaper job ads range between 0.59 (in 1950) to 0.81 (in 1990). 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, 2018's data set of online job postings also exhibited a similar under-representation of blue-collar occupations.)[37] Figure 10 presents the same set of relationships, now using a 4-digit SOC classification. Here, the correlations among the two measures of occupational size are weaker, ranging between 0.31 in 1960 to 0.48 in 1980. Figure 11 reproduces Figure 10 with occupations' census employment shares computed using only counts of workers from the Boston and New York MSAs.[38] The average correlation in the six panels of Figure 11 is 0.55, approximately 14 percentage points higher than in Figure 10.

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.[39] In Figure 12, we compare the undergraduate requirements across 4-digit SOC codes. Within

[37] While blue-collar workers are under-represented in our newspaper data relative to their employment shares, we emphasize that our analysis of changes in occupations' task content (or of economy-wide task content) is not affected by this under-representation, since we weight occupations by their employment shares.

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

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

Figure 9: Occupation Shares: Newspaper Vacancies versus U.S. Employment

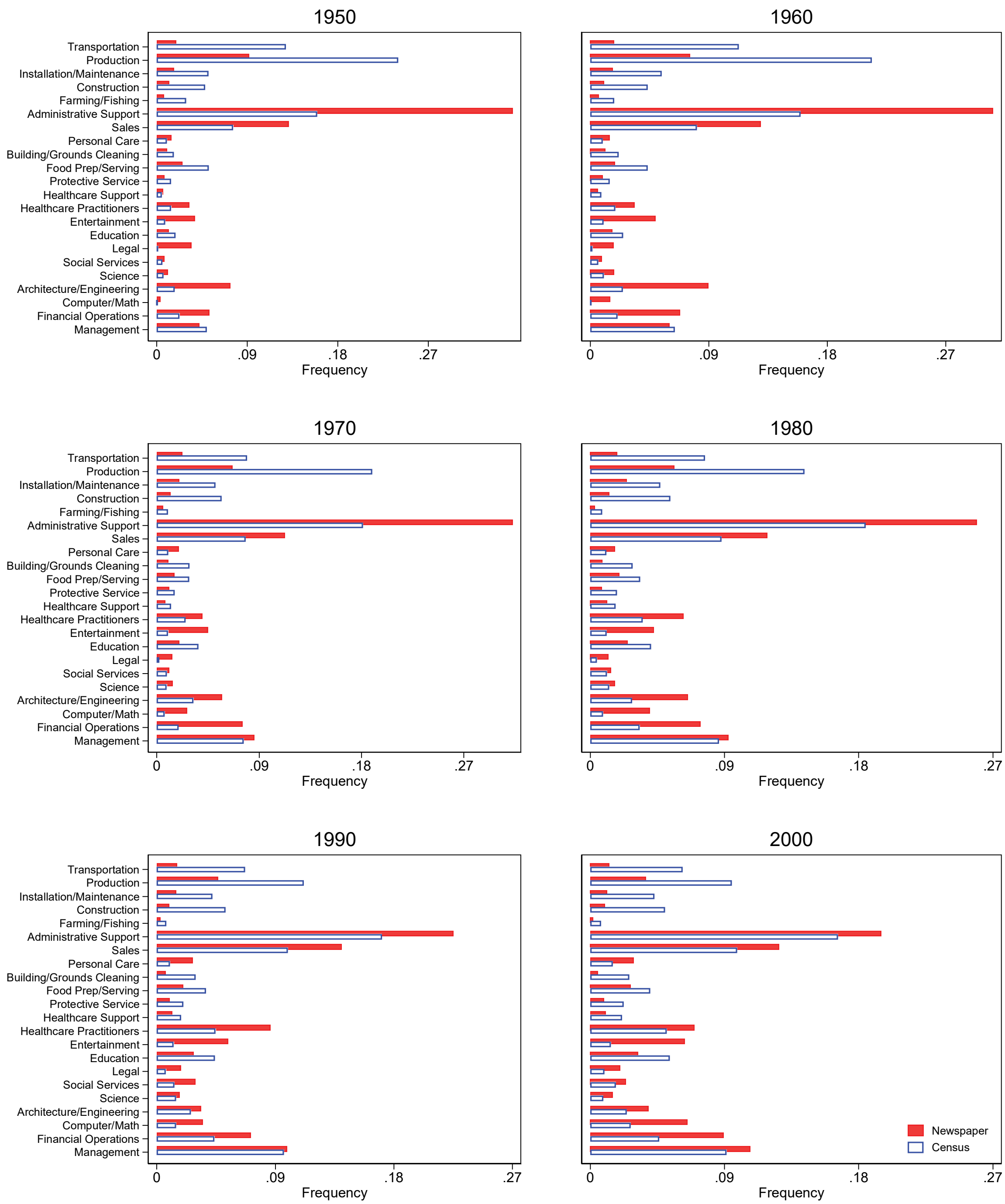
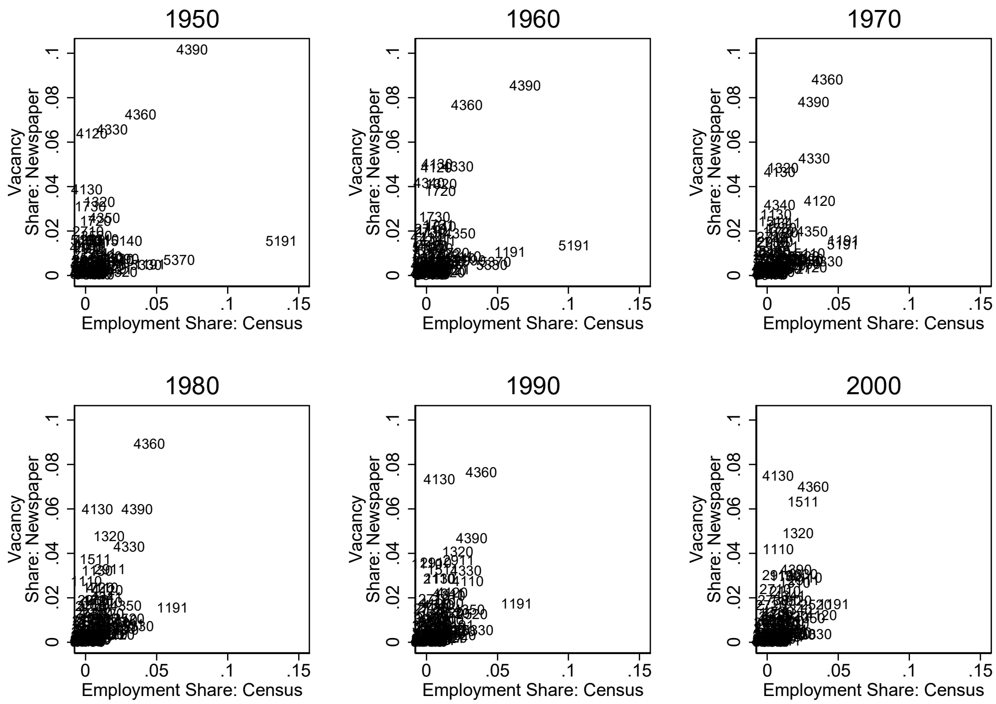


Figure 10: Occupation Shares: Newspaper Vacancies versus U.S. Employment



the individual years that are plotted, the correlations between the two measures are 0.46, 0.70, 0.66, 0.64, 0.63, and 0.62.

In Figure 13, 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 0.26. For the individual years in our sample, the correlations across SOC codes are 0.16, 0.18, 0.22, 0.32, 0.32, and 0.20. Overall, we conclude that the undergraduate degree requirement data which we extract from our newspaper data are correlated with the more cleanly measured census data on workers' educational attainment, but only weakly so for professional and graduate degrees.

B.2 Comparison to O*NET

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 words within the O*NET Element Description. We then append to our initial lists of words and phrases synonyms from our continuous bag of words model.

Figure 14 relates the O*NET Importance measure to newspaper keyword frequencies. Each panel 14 presents a comparison for a different O*NET Element: one work style (Cooperation),

Figure 11: Occupation Shares: Newspaper Vacancies versus Boston and New York MSA Employment

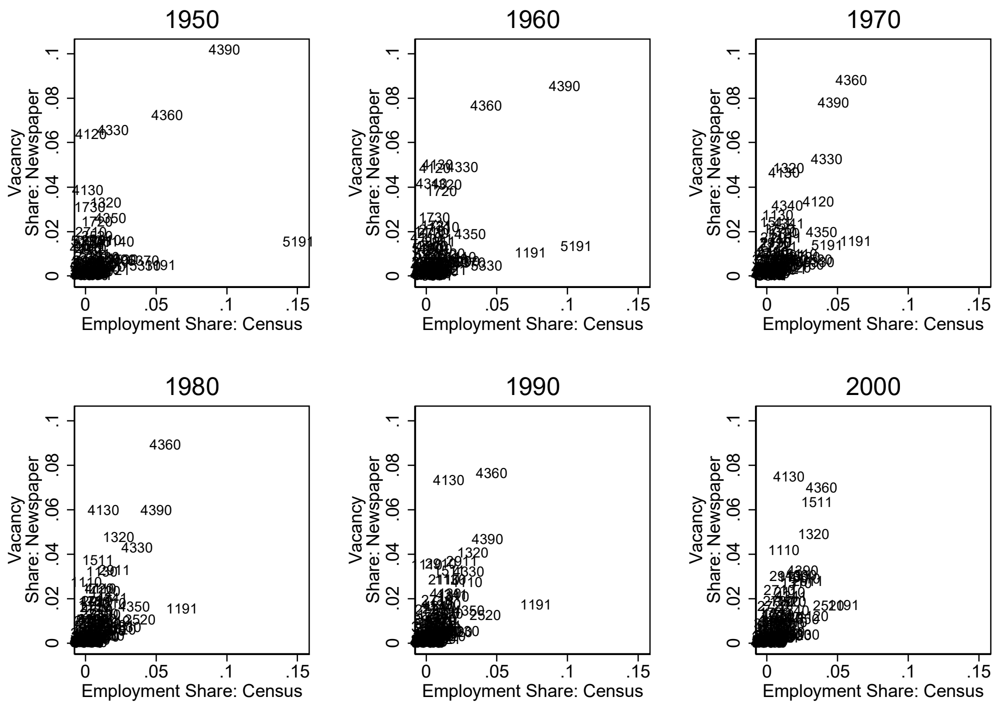
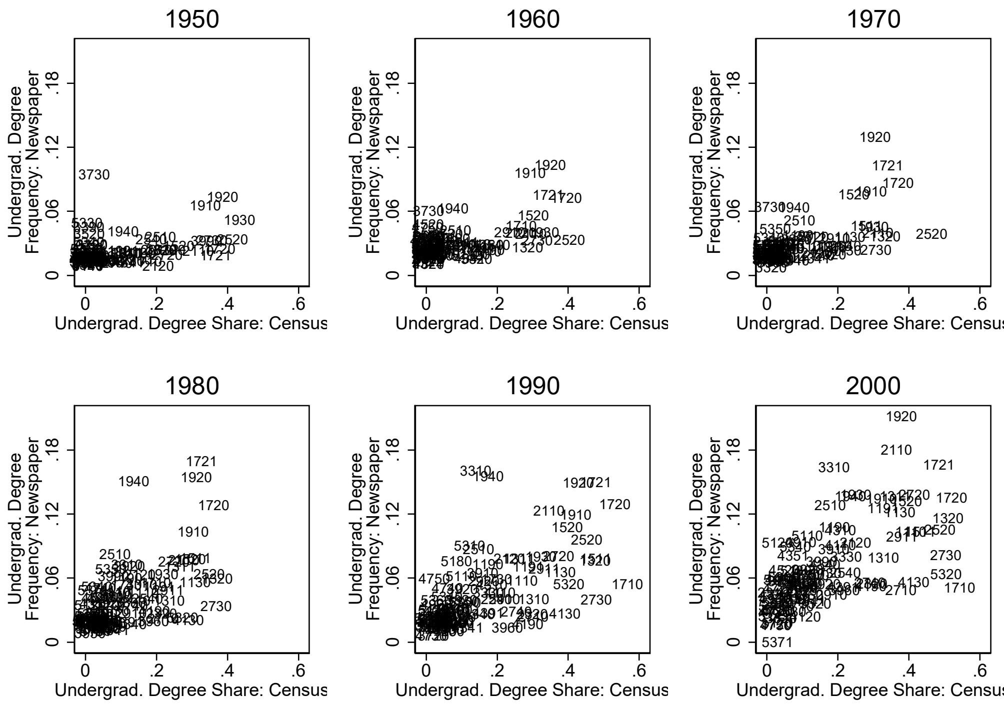


Figure 12: Educational Characteristics by Occupation



Notes: Each panel describes the relationship between the share of workers in each occupation with an undergraduate degree (according to the decennial census) on the x-axis; the fraction of newspaper ads which mention an undergraduate degree is on the y-axis.

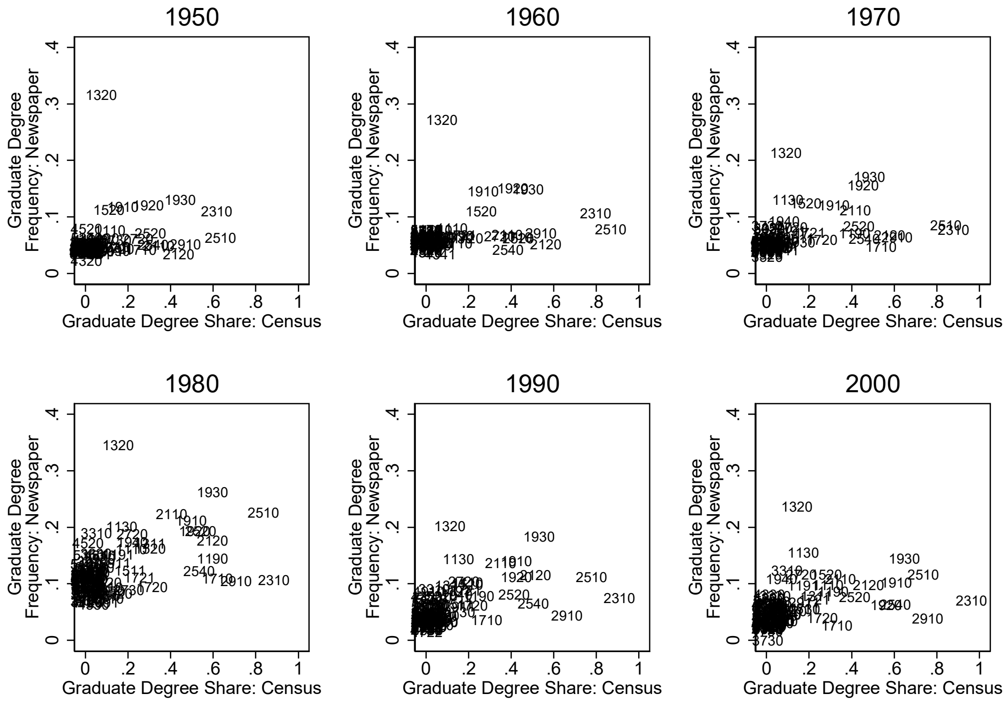
one skill (Active Listening), one knowledge requirement (Personnel and Human Resources), and one work activity (Operating Vehicles, Mechanized Devices, or Equipment). Within each panel, each data point represents a single 4-digit SOC occupation: For example, the "2310" (the code for Lawyers, Judges, and Related Workers) in the top right panel indicates that the O*NET Importance of the skill of "Active Listening" is 4.5 (on a scale from 1 to 5), while in the newspaper data, we detect 0.15 Management of Material Resources related keywords per 1000 (correctly spelled) job ad words. The correlations (weighted by the number of vacancy postings in our newspaper data) in these four plots are 0.27, 0.54, 0.54, and 0.37, respectively.[40]

The four relationships depicted in Figure 14 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 0.65 range, and are somewhat higher for knowledge requirements, skills, and activities (where the mean correlations are 0.57, 0.54, and 0.49, respectively) than for work styles (where the mean correlation is 0.32).[41]

[40] Across all 125 O*NET Elements, the unweighted correlations are lower by 3 percentage points on average.

[41] As we discuss in footnote 19, the O*NET database may not be the ideal benchmark for comparison, given its

Figure 13: Educational Characteristics by Occupation



Notes: Each panel describes the relationship between the share of workers in each occupation with a graduate degree (according to the decennial census) on the x-axis; the fraction of newspaper ads which mention a graduate degree is on the y-axis.

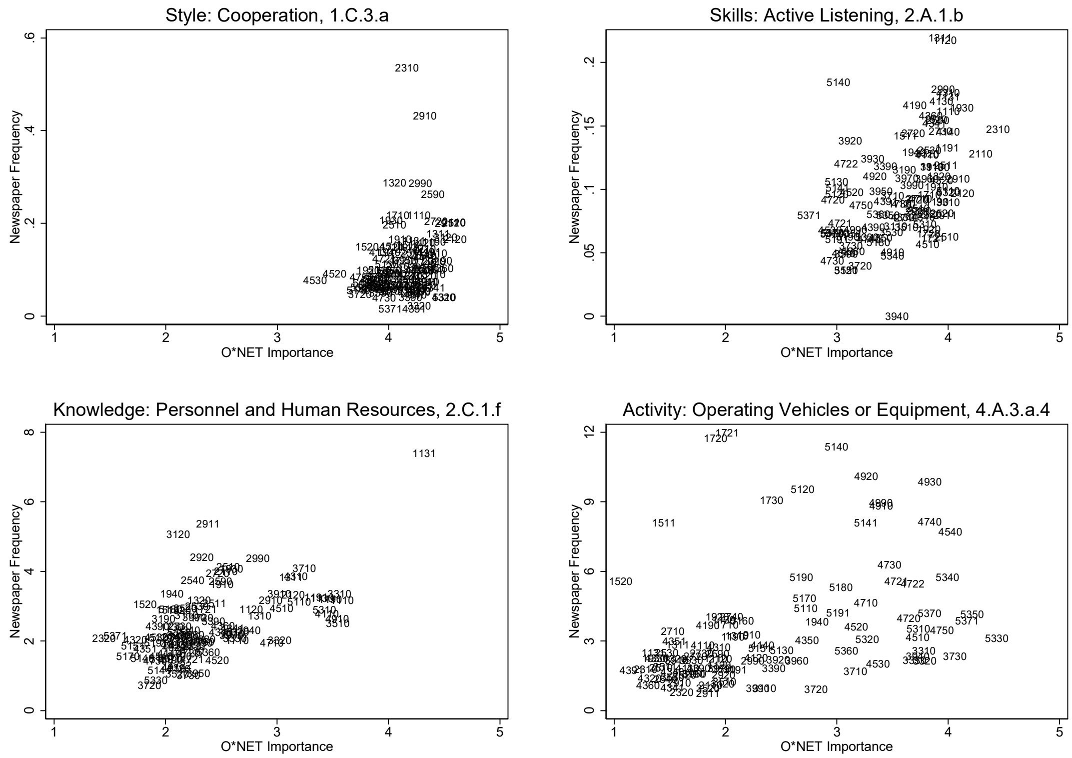
B.3 Comparison to the DOT

In Section II.C, we replicated Figure 1 of Autor, Levy, and Murnane (2003), one of the key empirical findings in the task literature. We showed that our new data set aligns with the Dictionary of Occupational Titles in its depiction of between-occupation shifts in task content. In this appendix, we provide an additional set of comparisons between our new data set and the Dictionary of Occupational Titles.

We first directly compare our newspaper measures to the analogous measures in a single vintage of the DOT. In the left panel of Figure 15, we plot the relationship between the Dictionary of Occupational Titles GED math measure, Autor, Levy, and Murnane (2003)'s benchmark measure of nonroutine analytic task intensity, and our newspaper-based nonroutine analytic task measure. From both data sets, we take values from 1977. The correlation between the two measures is 0.79. According to both data sets, engineering and computer-related occupations are those that have the highest nonroutine analytic task content. In the right panel of Figure 15, we present the analogous relationship for nonroutine interactive task measures. Here, the correlation is 0.16. The correlations for the other three measures are

well-known limitations for measuring occupational tasks (Autor, 2013). Nevertheless, we interpret these correlations as evidence that the newspaper text has valuable information about occupational tasks.

Figure 14: O*NET Importance Measures and Newspaper Keyword Frequencies



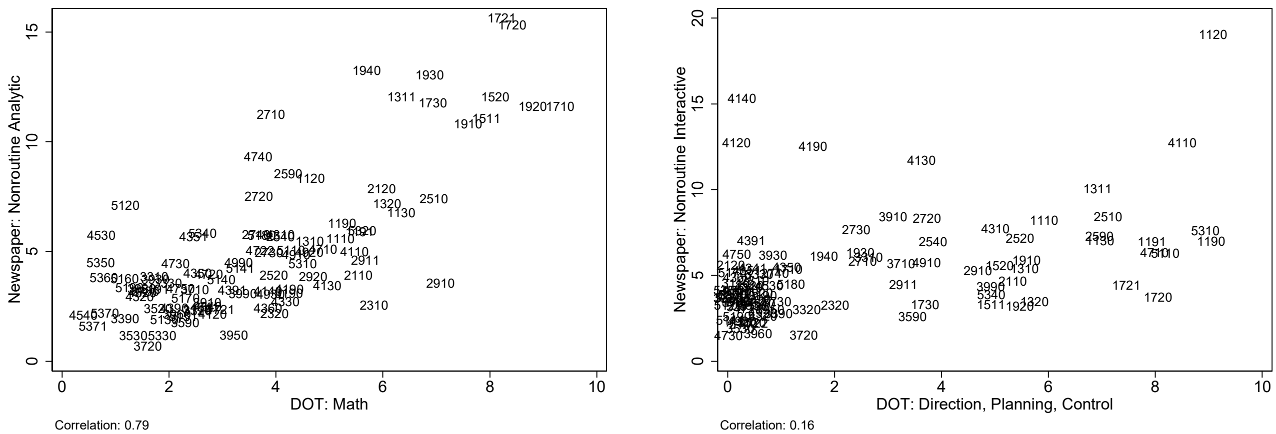
Notes: Each panel corresponds to one O*NET element. In each panel, each point represents a SOC code. The value of the x-axis represents the O*NET Importance measure (from version 22.1 of the O*NET database). The y-axis measures the number of keyword appearances per 1000 job ad words, using data from the *Boston Globe*, *New York Times*, and *Wall Street Journal*.

0.39 for nonroutine manual tasks, 0.34 for routine cognitive tasks, and 0.02 for routine manual tasks. So, for four out the five task groups, our newspaper-based task measures align at least moderately with those in the DOT.[42]

As far as we are aware, the 1977 and 1991 Dictionary of Occupational Titles (DOT) — which are the underlying source of the Autor, Levy, and Murnane measures — are the sole data set from which one could potentially measure within-occupation changes in U.S. task content over our sample period. We now present evidence that the DOT is ill-suited to measurement of within-occupation changes in tasks, corroborating the characterization made by Autor, Levy, and Murnane (2003), who note the "limited sampling of occupations (particularly in the service

[42] While we view DOT and O*NET as useful benchmarks, they too have issues, which extend beyond their limited ability to track within-occupation changes in job characteristics over time. In their review of the design of the O*NET data collection program, the National Research Council (2010) identifies several aspects of O*NET which may limit its usefulness as a research tool. Summarizing these issues, Autor (2013) writes that in both the DOT 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 DOT or O*NET measures, but we interpret the correlations as reassuring evidence that the newspaper text is a valuable source of task data.

Figure 15: Comparison of DOT and Newspaper Task Measures: 1977 Levels



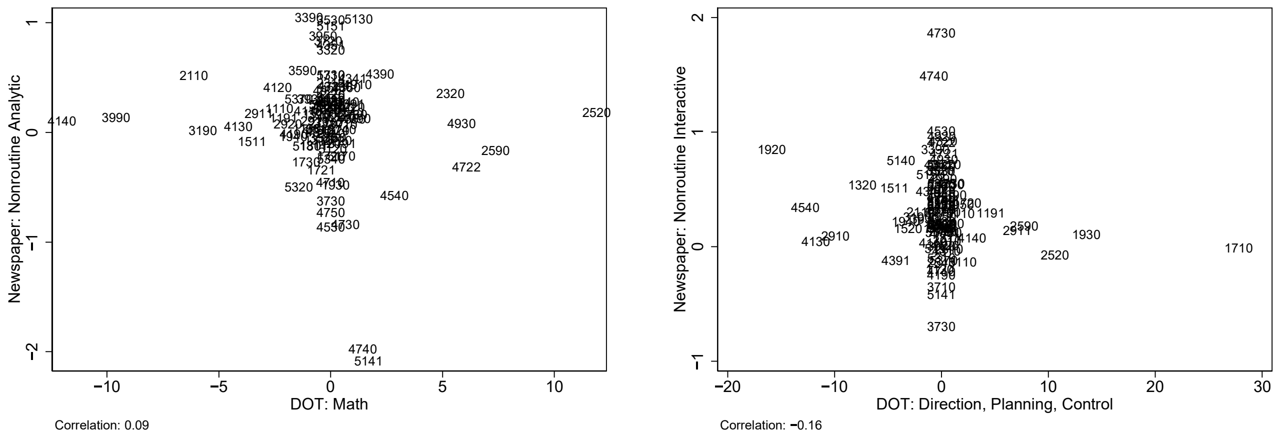
Notes: The left panel gives the relationship between the DOT GED math variable with our newspaper-based nonroutine analytic task measure (stated as number of mentions per 1000 job ad words). The right panel presents the same relationship for the DOT direction, planning, and control measure and our newspaper-based nonroutine interactive task measure. The numbers within the scatter plot characterize the SOC code. The stated correlation is computed with weights given by the newspaper number of job ads per occupation as of 1977.

sector), imprecise definitions of measured constructs, and omission of important job skills" (p. 1292-1293) limit the usefulness of the DOT for time series analysis.

While there is a substantial correlation in levels between our newspaper-based task measures and the DOT measures, no such correlation exists when looking at task growth rates. According to the DOT, from 1977 to 1991 there was a decline in nonroutine analytic tasks within occupations; routine manual tasks increased within occupations (see the bottom row of Table 6 of Autor, Levy, and Murnane, 2003). This is the opposite of what we find. Moreover, the correlation in occupations' changes in task intensity, from 1977 to 1991, is much weaker when comparing the two data sets. Figure 16 plots the correspondence for nonroutine analytic (left panel) and nonroutine interactive (right panel) tasks. For the five task measures, the correlations in growth rates in task intensities are: 0.09 (for nonroutine analytic tasks), -0.16 (nonroutine interactive), -0.11 (nonroutine manual), -0.04 (routine cognitive), and 0.14 (routine manual). Overall, there is essentially no relationship between the growth rates of DOT task measures and the growth rates of our newspaper-based task measures.

For many occupations, the DOT's measures were not updated between the 1977 and 1991 vintages. Suggestive of this, the correlation across occupations in the GED math scores across the 1977 and 1991 vintages of the DOT equals 0.98. To highlight this serial correlation and illustrate the limited nature of the time series that is available with the DOT, we plot the task measures in the 1977 and 1991 editions, in Figure 17. This figure indicates that, for a large fraction of occupations, the GED math measure (left panel) and the direction, planning, and control measure (right panel) are essentially the same across DOT editions. The correlations for the three un-plotted tasks are 0.95 for finger dexterity; 0.95 for eye, hand, and foot co-

Figure 16: Comparison of DOT and Newspaper Task Measures: 1977 to 1991 Growth Rates



Notes: The left panel gives the relationship between the DOT GED math variable with our newspaper-based nonroutine analytic task measure (stated as a growth rate between 1977 and 1991). The right panel presents the same relationship for the DOT direction, planning, and control measure and our newspaper-based nonroutine interactive task measure. The numbers within the scatter plot depict the SOC Code. The stated correlation is computed with weights given by the number of newspaper job ads per occupation as of 1977 and 1991.

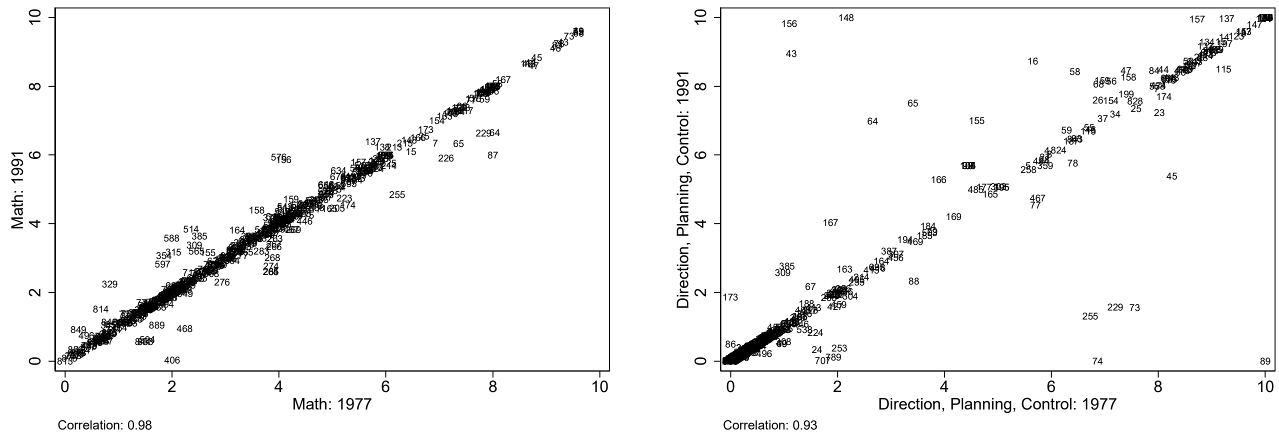
ordination; and 0.87 for setting tolerances.[43] These correlations, especially those for the GED math variable, are suggestive of irregular and incomplete updating of occupations' task content measures between 1977 and 1991.

It is hypothetically possible that there were actually no task changes, within occupations, for a large number of occupations. This point of view, however, is inconsistent National Research Council (1999) and Spitz-Oener (2006). To be clear, for the main empirical exercise that Autor, Levy, and Murnane (2003) perform—measuring the relationship between occupations' task content changes and changes in computer adoption rates—the fact that many task measures were not updated by DOT examiners does not pose a problem. However, these plots do indicate that the DOT is ill-suited in measuring within-occupation changes in task content for a broad swath of occupations.

In summary, across occupations, our task measures broadly align with those in the Dictionary of Occupational Titles. However, the portrayals of within-occupation task trends — comparing our newspaper data with the DOT—are starkly different. In contrast to what our newspaper data indicate, the DOT data set indicates that there has been a shift within occupations away from nonroutine analytic tasks towards routine manual tasks.

[43] The unweighted correlations are higher than the weighted correlations by 0.02, on average.

Figure 17: Comparison of 1977 and 1991 DOT



Notes: The left panel gives the relationship, according to the 1977 and 1991 editions of the DOT, of occupations' GED math variable. The right panel presents the same relationship for the DOT direction, planning, and control measure. The numbers within the scatter plot are the 1980-90 occupation code, as defined by Autor, Levy, and Murnane. The stated correlation is computed with employment weights, given by summing across individuals working in each occupation as sampled in the 1984 CPS.

B.4 Top Occupations

In addition to the measures developed by Spitz-Oener (2006), we apply the mapping between keywords and skills that Deming and Kahn (2018) use 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 16, 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 (2018). As one would expect, ads for sales representatives and sales managers have the highest frequency of words related to customer service skills; ads for financial specialists and accountants have the highest frequency of words related to financial skills; ads for health diagnosticians have the highest frequency of words related to people management; and ads for engineers have the highest frequency of words related to project management.

C Selection and Measurement Error

C.1 Trends in Ad Length and Spelling Accuracy

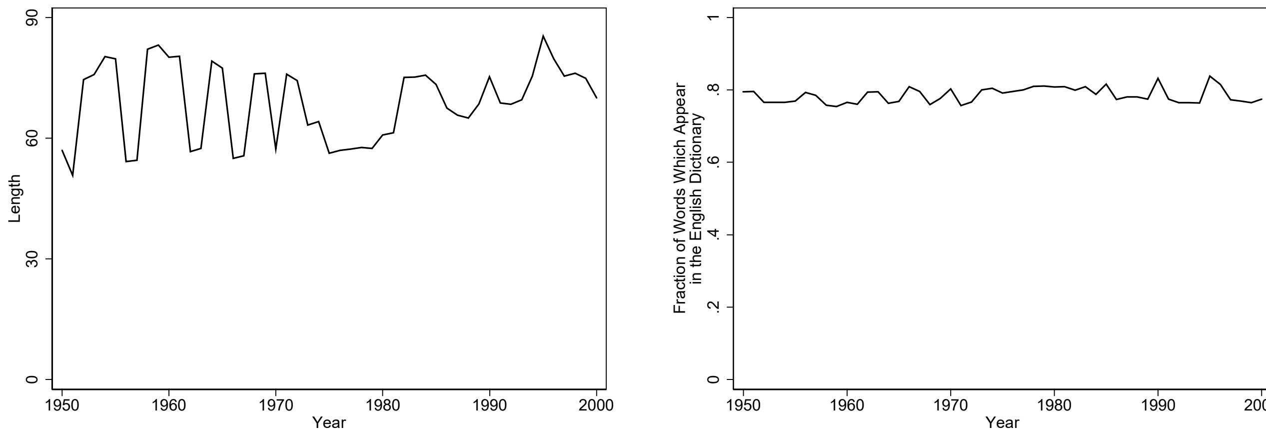
In Figure 18, we plot the average length of each ad, and the fraction of correctly spelled words for each ad. These plots indicate that there is a weak upward trend in ad length over our sample period — the average number of words per ad in our data set is 69.2 in the 1950s, 69.3 in the 1960s, 62.0 in the 1970s, 68.8 in the 1980s, and 74.5 from 1990 to 2000 — and that there is no trend in the fraction of words that are correctly spelled (i.e., words that are in our English dictionary). The motivation for these plots is that any time-varying measurement

Table 8: Top Occupations According to the Deming and Kahn (2018) Classification of Skills

Character			Computer		
Retail Sales	0.0004	13.23	Systems Analyst	0.0010	18.41
Dental Assistant	0.0007	12.27	Systems Engineer	0.0004	17.50
Management Trainee	0.0012	11.58	Programmer Analyst	0.0014	15.35
Real Estate Sales	0.0007	11.34	Computer Operator	0.0005	13.41
Secretary Administrative Assistant	0.0005	10.73	Programmer	0.0035	13.02
Customer Service			Financial		
Retail Sales	0.0004	17.41	Financial Analyst	0.0009	26.00
Sales Representative	0.0017	15.07	Staff Accountant	0.0004	24.87
Sales Manager	0.0018	14.71	Assistant Controller	0.0007	22.73
Sales Engineer	0.0012	14.12	Accountant Junior	0.0009	22.26
Sales Trainee	0.0011	13.95	Accountant	0.0010	22.06
People Management			Problem Solving		
Executive Director	0.0007	9.67	Physicist	0.0004	7.34
Management Trainee	0.0012	7.58	Chemist	0.0017	7.04
RN	0.0070	7.34	Statistical Typist	0.0009	6.72
Physical Therapist	0.0007	7.05	Chemical Engineer	0.0003	4.90
Director	0.0033	6.87	Systems Engineer	0.0004	4.34
Project Management			Social		
Project Engineer	0.0006	28.96	Social Worker	0.0015	6.97
Project Manager	0.0008	24.59	Executive Director	0.0007	4.10
Mechanical Engineer	0.0010	24.05	Worker	0.0005	4.03
Design Engineer	0.0005	22.48	Systems Engineer	0.0004	3.91
Electrical Engineer	0.0007	21.85	Account Executive	0.0010	3.19
Writing					
Editorial Assistant	0.0005	8.45			
Editor	0.0022	8.27			
Technical Writer	0.0007	7.97			
Writer	0.0010	6.83			
Proofreader	0.0009	3.19			

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 $1/51 \cdot \sum_{t=1950}^{2000} S_{jt}$ — the average share of ads belonging to the job title; and the final column gives the frequency of task h words among job title j's ads, per 1000 ad words. Footnote 16 contains the words and phrases corresponding to each skill that were used in Deming and Kahn (2018). To these lists, we append similar words, using the continuous bag of words model introduced in Section I.C.

Figure 18: Trends in Ad Length, Fraction of Correctly Spelled Words



Notes: The left panel plots the average length (including both words that appear in the English dictionary and those that do not). The right panel plots the fraction of words within each ad that are English-dictionary words. The correlation between year and ad length is 0.27 (with a p-value of 0.088), and between year and the fraction of words which are correctly spelled is 0.09.

error in our newspaper text would manifest as trends in the share of correctly spelled words. Reassuringly, no such trend is apparent.

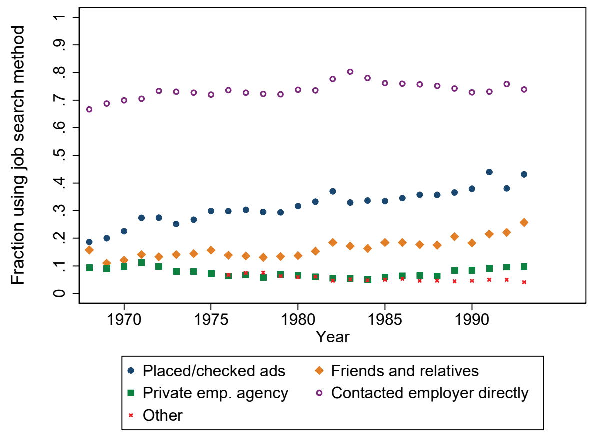
C.2 Methods of Job Search

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 unemployed civilian workers, who are looking for a job, and who are between the ages of 16 and 64. Figure 19 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 19 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,

Figure 19: 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.

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.

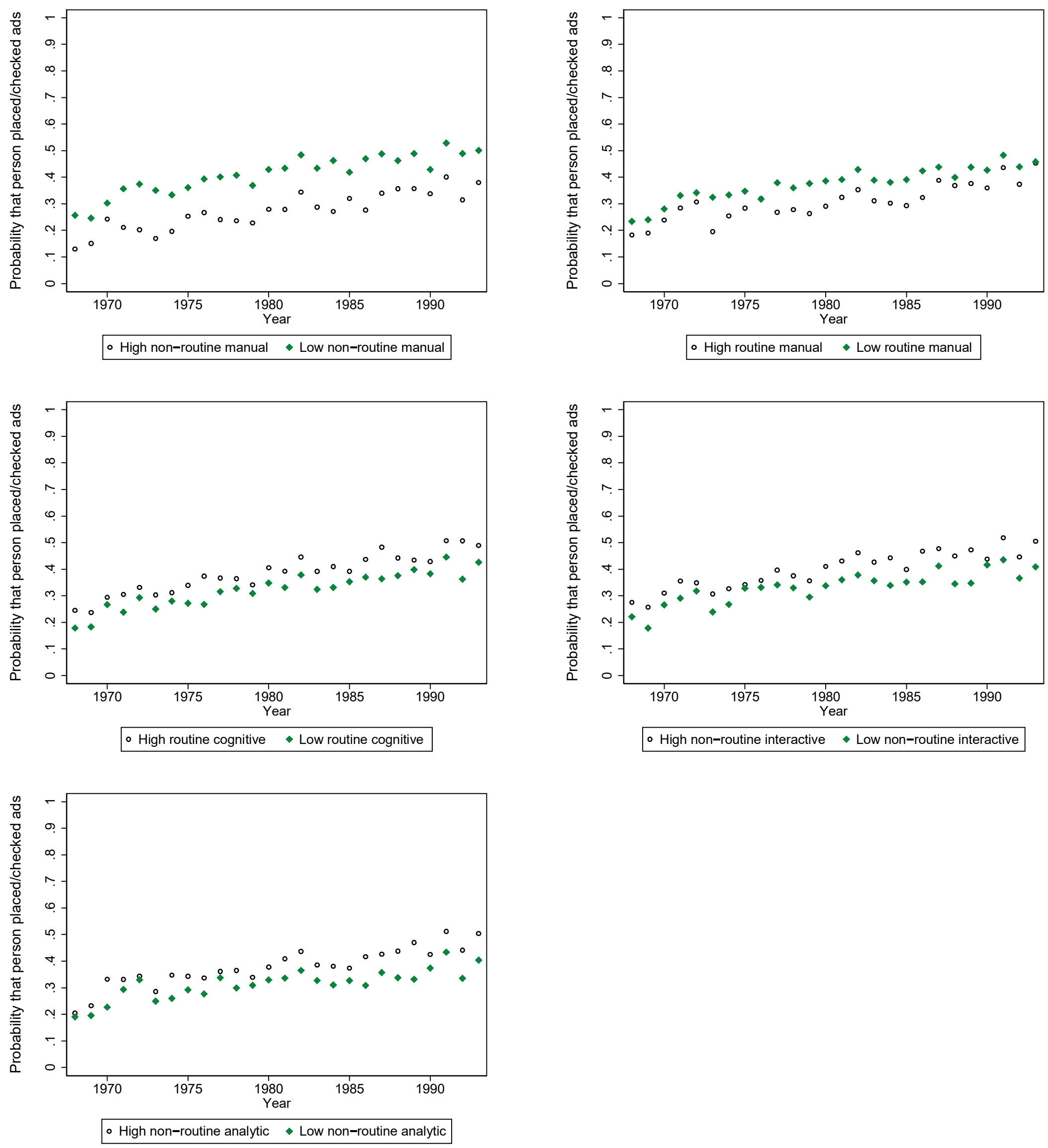
To test this hypothesis, we first compute the mean task content in each occupation over the entire sample period. We then plot the fraction of workers searching for jobs through ads whose last occupation was highly intensive in, say, nonroutine interactive tasks (75th percentile or higher) and the same for workers in occupations that have a low intensity in the same task (25th percentile or lower).[44]

Figure 20 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.[45]

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

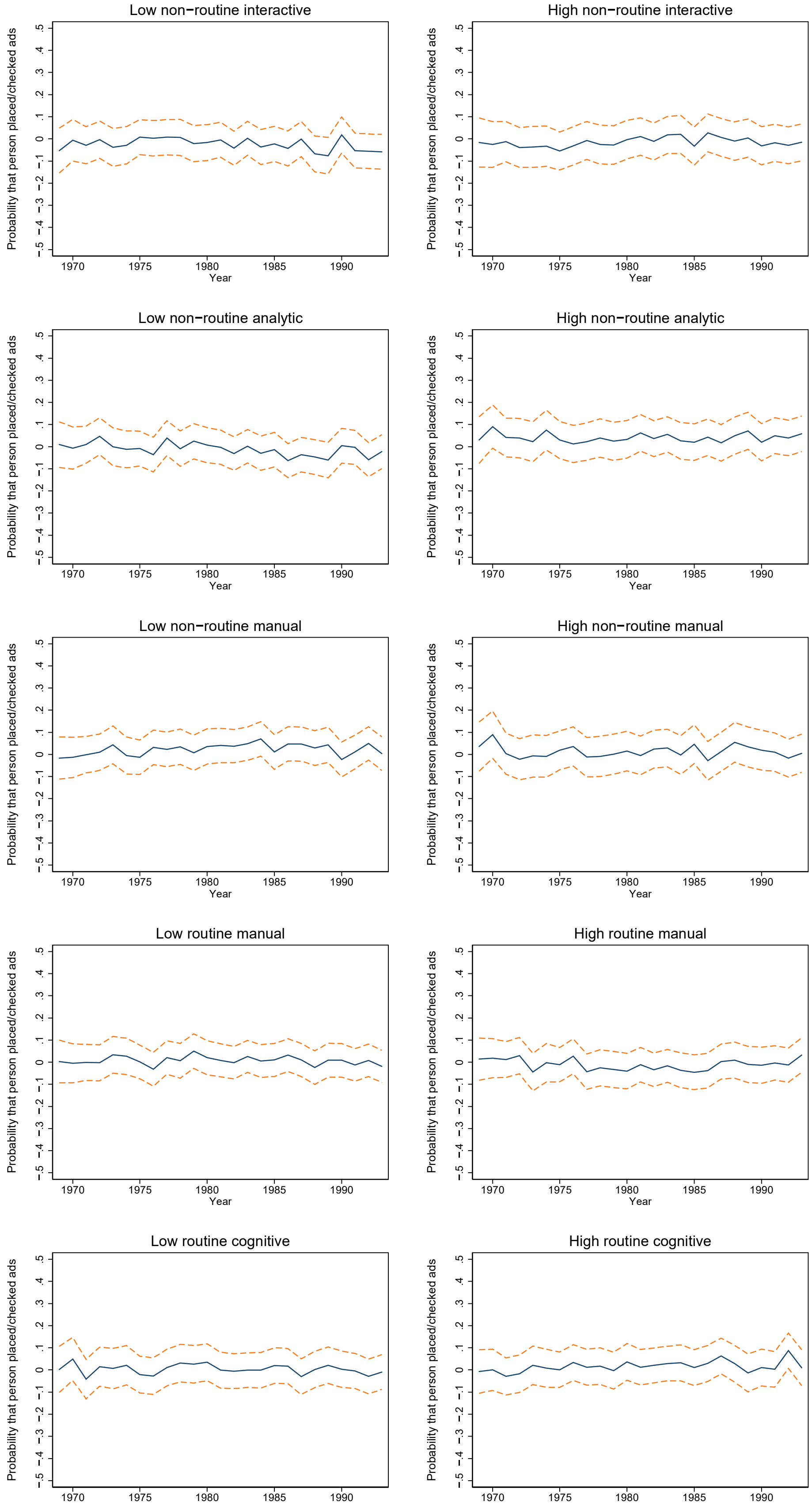
[45]While trends in job search do not appear to differ by occupation, there are *level* differences in job search: For example, workers in high routine cognitive occupations are more likely to use job ads as a search method compared with workers in low routine cognitive occupations. This pattern is consistent with our finding in Appendix B 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

Figure 20: 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 21: Trends in Job Search Method by Task Intensity of Occupation



Notes: The figure above plots the estimates for β_t in Equation 3 for each of the five tasks, along with the 95 percent confidence intervals. Each panel represents the results of a separate regression.

We test this hypothesis formally using the following regression:

$$y_i = \beta_0 + \beta_t \widetilde{T}_h^\tau + x_i^{'}\gamma + \iota_t + \epsilon_i, \tag{3}$$

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). 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 21 plots the estimates for β_t. The omitted year is 1968, so coefficients β_t are interpreted as relative to 1968. Overall, Figure 21 suggests no detectable trends in job search behavior through ads.

C.3 Representativeness of Boston and New York Job Ads

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. Workers in these metro areas tend to have higher education, are paid higher wages, and are over-represented in certain types of occupations (e.g., in financial management, in tertiary education, etc.) and under-represented in others. What is more, this non-representativeness may be growing over time (for example, the college graduate share in New York City and Boston has increased faster than in other parts of the country.)

Unfortunately, we cannot examine — based on our newspaper data — whether our occupations' task measures are substantially different in Boston and New York compared to the rest of the United States. However, we have a sample of text from vacancy postings from a more recent period, from October 2011 to March 2017, from which we can examine the representativeness of Boston and New York. Our sample is drawn from a 5 percent sample of ads, 7.6 million ads, which were collected by EMSI.[46]

To do so, we begin by computing our nonroutine and routine task measures, using the same mapping between words and task groups that we use in the rest of the paper. We then perform a set of regressions characterized by the following equation:

$$\text{task}_{ajt}^h \quad = \quad \beta_h \cdot 1_{a \in \{\text{Boston, New York}\}} + \iota_{jh} + \iota_{th} + \iota_{sh} + \epsilon_{ahjt}. \tag{4}$$

In Equation 4, h refers to one of five routine and nonroutine task categories; task_{ajt}^h equals the number of mentions of task h (relative to the number of words in the ad) in a, published

side search behavior.

[46]For this exercise, we drop the first three months — October 2011 to December 2011 — as the number of ads collected per month is rapidly expanding over the very beginning of the EMSI sample period (suggesting that, relative to the rest of the sample period, the samples in the first few months may not be representative.)

Table 9: Estimates from Equation 4

Fixed Effect	Nonroutine Analytic	Nonroutine Interactive	Nonroutine Manual	Routine Cognitive	Routine Manual
None	0.341	0.319	-0.125	-0.012	0.046
	(0.011)	(0.011)	(0.004)	(0.003)	(0.003)
4-digit SOC	0.142	0.168	-0.085	-0.019	0.061
	(0.010)	(0.010)	(0.004)	(0.003)	(0.003)
6-digit SOC	0.074	0.177	-0.083	-0.025	0.052
	(0.010)	(0.010)	(0.004)	(0.003)	(0.003)
Job Title	-0.023	0.078	-0.062	-0.014	0.024
	(0.009)	(0.009)	(0.004)	(0.003)	(0.002)
Mean of h	5.246	5.028	1.019	0.596	0.250
Std. Dev. of h	6.330	6.236	2.401	1.929	1.547

Notes: Each column presents the coefficient estimates, standard errors, and sample statistics of one of our five Spitz-Oener task measures. The first four set of rows present coefficient estimates and standard errors of β_h, with each set of rows applying a different occupation fixed effect. The sample in these regressions includes the job ads for which we could retrieve an SOC code based on the ad's job title, 5.5 million job ads. The final two rows present the average and standard deviation of the task measure in our sample of 5.5 million ads.

in year t, for an occupation j; ι_{jh}, ι_{th}, and ι_{sh} respectively refer to occupation fixed effects, year fixed effects, and fixed effects for the job message board from which EMSI procured the data.[47] The coefficient of interest is β_h, characterizing the relative frequency of task h in the Boston and New York metro areas, relative to the rest of the U.S. We conduct regressions with increasingly detailed occupation fixed effects. These regressions will allow us to assess the extent to which, for example, Engineers in New York and Boston differ from those in the rest of the U.S. (using a 4-digit SOC fixed effect), Electrical Engineers in New York and Boston are unique (using a 6-digit SOC fixed effects), or Wire Design Installation Engineers in New York and Boston are unique (using job title fixed effects.)

The results from this exercise are given in Table 9. We find substantial differences in the overall task content of Boston and New York jobs, relative to the rest of the country. Per 1000 job ad words, there are 0.34 additional nonroutine analytic task words (0.054= 0.34/6.33 standard deviations) and 0.32 nonroutine interactive task words (0.051 standard deviations) in Boston and New York. However, much of the differences are due to the fact that the occupational mix of Boston and New York are different from that of the country as a whole (as opposed to individual occupations differing in their task content). Using 4-digit SOC fixed effects, the nonroutine analytic and interactive task content of jobs in New York and Boston are higher by 0.022 and 0.027 standard deviations, respectively. Using more detailed fixed effects, at the level of 6-digit SOC codes or job titles, leads to even smaller discrepancies between our two metro areas and the rest of the U.S.

We are not only specifically interested in the level of non-representativeness of our New

[47] As the sample period has progressed, EMSI has collected job ad text from an increasingly wide variety of job posting websites. We include website-specific fixed effects to account for the changing composition over the period.

York and Boston newspaper text, but also in trends in non-representativeness over our 1950 to 2000 sample period. In fact, since our contribution relates to within-occupation trends in task content, trends in non-representativeness of New York and Boston would be especially problematic. For the short (five-year) period from which we have online job ad text, we can examine whether there are any trends in the task content of New York and Boston jobs relative to jobs in other portions of the U.S. To this end, we examine regressions characterized by the following equation:

$$\text{task}_{ajt}^{h} = [\beta_h + \gamma_h \cdot (t - 2012)] \cdot 1_{a \in \{\text{Boston, New York}\}} + \iota_{jh} + \iota_{th} + \iota_{sh} + \epsilon_{ahjt} \qquad (5)$$

The parameter of interest in Equation 5 is γ_h; it characterizes the growth rate in task h mentions over the sample period in Boston and New York compared to the rest of the U.S. As before, inclusion of occupation fixed effects tends to reduce the magnitude of the γ_h coefficients. The one coefficient estimate of γ_h that is most indicative of substantial trends in the non-representativeness of New York and Boston jobs is that of routine manual tasks with job title, 4-digit SOC, or 6- digit SOC fixed effects. Using job title fixed effects, ads from Boston and New York (relative to the rest of the United States) mentioned an additional 0.04 (=0.007·5) routine manual task words (per 1000 job ad words) in 2017 relative to 2012. This difference represents 0.023 standard deviations of the routine manual task measure. For the other four task measures, the trend in the difference between Boston and New York task mentions and task mentions in the rest of the country is at most 0.020 standard deviations of the task measure.

Overall, while there are statistically significant differences in ads posted in Boston and New York, relative to the U.S., these differences largely exist between occupations rather than within occupations. Moreover, the differences between New York and Boston and the rest of the U.S. are modest, when compared to the overall dispersion in task measures, across all online job ads.

D Details on the Construction of the Database

This section provides further details that, due to space constraints, we could not include in Section I. 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 I.

Table 10: Estimates from Equation 5

Fixed Effect		Nonroutine Analytic	Nonroutine Interactive	Nonroutine Manual	Routine Cognitive	Routine Manual
None	β_h	0.375	0.275	-0.172	-0.027	-0.026
		(0.025)	(0.025)	(0.010)	(0.008)	(0.006)
	γ_h	0.034	0.032	-0.018	-0.005	0.001
		(0.005)	(0.005)	(0.002)	(0.002)	(0.001)
4-digit SOC	β_h	0.084	0.143	-0.088	-0.030	-0.001
		(0.022)	(0.023)	(0.009)	(0.007)	(0.006)
	γ_h	-0.006	0.001	-0.009	-0.001	0.001
		(0.004)	(0.004)	(0.002)	(0.001)	(0.001)
6-digit SOC	β_h	-0.059	0.079	0.083	0.026	0.128
		(0.039)	(0.039)	(0.015)	(0.012)	(0.010)
	γ_h	-0.009	0.003	0.002	0.001	0.020
		(0.007)	(0.007)	(0.003)	(0.002)	(0.002)
Job Title	β_h	-0.017	0.061	0.010	0.009	0.092
		(0.035)	(0.036)	(0.014)	(0.012)	(0.009)
	γ_h	0.002	0.025	-0.005	-0.003	0.007
		(0.006)	(0.006)	(0.002)	(0.002)	(0.002)
Mean of h		5.246	5.029	1.020	0.596	0.250
Std. Dev. of h		6.330	6.236	2.401	1.929	1.547

Notes: Each column presents the coefficient estimates, standard errors, and sample statistics of one of our five Spitz-Oener task measures. The first four sets of rows present coefficient estimates and standard errors of β_h and γ_h, with each set of rows applying a different occupation fixed effect. The final two rows present the average and standard deviation of the task measure in our sample of 5.5 million ads.

D.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 words can be found. This solution, however, could be problematic as, for example, the word “sales” appears in vacancy postings for “sales representatives” and to advertise retail sales. 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.

In our context, the Latent Dirichlet Allocation (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 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 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." Using this intuition, we apply 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 et al. (2003, pp. 996-998).

Notation and Terminology

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

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

   where each $\alpha_k > 0$ and $\Gamma(\cdot)$ 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 $\beta = \Pr(w_n = 1 | z_k = 1)$.

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:

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

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:

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

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 et al. (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 100 thousand pages of advertisements from 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 cutoff 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.[48]

LDA Results

In Table 11, we present partial results from our LDA procedure, listing the ten words which are most predictive of each topic for each of the subsamples in our sample: 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. 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}.

D.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, cities, and address fragments, such as such as "150 East 42 Street, N. Y."

- 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.

- 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

[48]Within economics, we are aware of one application of LDA, which classifies Federal Open Market Committee statements by topic. See Fligstein et al. (2014) and Hansen et al. (2018).

Table 11: Predictive Word Stems in the LDA Model

Panel A: *Boston Globe* Classified Ads	
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 Ads	
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 Ads	
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 Ads	
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 Ads	
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

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' task contents 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, in the body of the paper, this would occur in the lines containing "ACCOUNTANTS," "MECHANICAL ENGINEER "METHODS ENGR," "RESUMES PRINTED," "Needs A," "TRANSPORTATION ADVERTISING SUPERVISOR," "With ...," "PERFORMANCE ENGINEERS", and "RESEARCH LABORATORIES UNITED AIRCRAFT CORPORATION." The lines with "ACCOUNTANTS," "MECHANICAL ENGINEER," "TRANSPORTATION ADVERTISING SUPERVISOR," and "PERFORMANCE ENGINEER" also contain a personal-noun job title. As a result, the contents of these lines are reported as the ads' job titles. Our algorithm correctly recognizes that "RESUMES PRINTED" and "Needs A" do not refer to job titles, but also erroneously omits "METHODS ENGR" from the list of job titles.

D.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 et al. (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 our 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.[49]

The objective, which we will try to maximize via maximum likelihood, is given by the probability of observing a word w_i 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 the MLE objective function, we omit words w_i which appear fewer than five times in our corpus.[50]

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 October 2011 and March 2017. As with our newspaper data, these text contain a job title for each vacancy posting. 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

[49] 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 $\overline{w}$.

[50] Mikolov et al. (2013a) estimated a model using text from a Google News corpus, and found that increasing the dimensionality of the model's vector representation from 300 to 600 led to only small improvements in performance.

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").

From this shorter list, we apply a *continuous bag of words* model in combination with an ancillary data set provided to us by EMSI (see Bengio et al., 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.

Manually retrieving SOC codes for all of the job titles in our data set would be infeasible. Among ads posted on or after 1950, there are, after all, more than 306 thousand unique job titles which are mentioned in at least two job ads, and more than 75 thousand unique job titles which are mentioned in at least five job ads. We retrieve SOC codes using our continuous bag of words model. In particular, for each job title $\mathcal{N}$ in our newspaper data, we compute the similarity between $\mathcal{N}$ and all of the job titles, $\mathcal{O}$, which appear in O*NET's (version 22.1) either Sample of Reported Titles or Alternate Sample of Reported Titles. For each O*NET job title $\mathcal{O}$, we observe an SOC code. So, for the job title $\mathcal{N}$, we assign to $\mathcal{N}$ the SOC code of the O*NET job title $\mathcal{O}$ closest to $\mathcal{N}$. We do this for any job title that appears our newspaper data.

In a second step, we assign an SOC code of 999999 ("missing") if certain words or phrases appear—"associate," "career builder," "liberal employee benefit," "many employee benefit," or "personnel"—anywhere in the job title, or for certain exact titles: "boys," "boys boys," "men boys girls," "men boys girls women," "men boys men," "people," "professional," or "trainee." These words and phrases appear commonly in our newspaper ads and do not refer to the SOC code which our CBOW model indicates. "Associate" commonly appears the part of the name of the firms which are placing the ad. "Personnel" commonly refers to the personnel department to which the applicant should contact.

We also replace the SOC code for the job title "Assistant" from 399021 (the SOC code for "Personal Care Aides") to 436014 (the SOC code for "Secretaries and Administrative Assistants"). "Assistant" is the fourth most common job title, and judging by the text within the job ads refers to a secretarial occupation rather than one for a personal care worker. While we are hesitant to modify our job title to SOC mapping in an ad hoc fashion for any job title, mis-specifying this mapping for such a common title would have a noticeably deleterious impact on our data set.

In a final step, we amend the output of the CBOW model for a few ambiguously defined job titles. These final amendments have no discernible impact on aggregate trends in task content, on role within-occupation shifts in accounting for aggregate task changes, or on the role of shifts in the demand for tasks in accounting for increased earnings inequality. First, for job titles which include "server" and which do not also include a food-service-related word — banquet, bartender, cashier, cocktail, cook, dining, food, or restaurant — we substitute an SOC code beginning with 3530 with the SOC code for computer systems analysts (151121). Second, for job titles which contain the word "programmer," do not include the words "cnc" or "machine," we substitute SOC codes beginning with 5140 or 5141 with the SOC code for computer programmers (151131). Finally, for job titles which contain the word "assembler" and do not contain a word referring to manufacturing assembly work — words containing the strings "electronic," "electric," "machin," "mechanical," "metal," and "wire" — we substitute SOC codes beginning with 5120 with the SOC code of computer programmers (151131). The amendments, which alter the SOC codes for approximately 0.2 percent of ads in our data set, are necessary for ongoing work in which we explore the role of new technologies in the labor market. Certain words refer both to a job title unrelated to new technologies as well as to new technologies. By linking the aforementioned job titles to SOCs that have no exposure to new technologies, we would be vastly overstating the rates at which food service staff or manufacturing production workers adopt new information and communication technologies. On the other hand, since these ads represent a small portion of the ads referring to computer programmer occupations, lumping the ambiguous job titles with the computer programmer SOC codes will only have a minor effect on the assessed technology adoption rates for computer programmers.

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 manual.* 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 that have a cosine similarity greater than 0.55 to any of the words in footnote 14. We also append any additional words that 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. First, with the aim of emulating O*NET's database, we construct our own classification 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 I.B, built from the newspaper 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.[51] On the other hand, the continuous bag of words model has the potential to identify words as synonyms 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, seems more of a stretch.

Our second classification scheme applies the mapping between keywords and skills which Deming and Kahn (2018) define in their study of the relationship between firms' characteristics and the skill requirements in their vacancy postings.[52] To each of these lists of words, we 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.

[51] 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. A second advantage of our CBOW model is that it allows us to partially undo the transcription errors generated in ProQuest's image scanning. Our CBOW algorithm, for example, identifies "adverhslng" as synonymous "advertising."

[52] See Table 1 of Deming and Kahn (2018) for their list of words and their associated skills. Building on their definitions, we use the following rules (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 management, staff, supervisory; (8) financial: accounting, budgeting, cost, finance, financial; (9) computer (general): computer, software, spreadsheets.

Table 12: Summary of Robustness Checks

Table Number	Measure	Occupation Classification	Description of Exercise	Avg. Within Share
3	Spitz-Oener	6-Digit SOC	Benchmark	0.879
4	Spitz-Oener	Job Title	Benchmark	0.883
5	Deming and Kahn	Job Title	Benchmark	0.861
13	Spitz-Oener	4-Digit SOC	Benchmark	0.944
14	Spitz-Oener	Job Title	No Employment Weights	0.786
15	Spitz-Oener	Job Title	Sample: NYT Classified Ads	0.888
16	Spitz-Oener	Job Title	Sample: NYT Display Ads	0.899
17	Spitz-Oener	Job Title	No Words from CBOW	0.969
18	Spitz-Oener	Job Title	Account for Employment Turnover	0.846
19	Spitz-Oener	Job Title	Normalized Task Measures	0.759
20	Deming and Kahn	Job Title	Normalized Task Measures	1.000

Notes: This table summarizes the average "Within Share" for various decomposition. Within this table, the second column gives the task measure used in the decomposition, either Spitz-Oener (2006)'s or Deming and Kahn (2018)'s categorization. The third column gives the level of occupational aggregation. The fifth column describes other dimensions along which we vary the sample, weighting, or normalization of our decompositions. We provide the formula for the "Avg. Within Share" footnote 22.

E Robustness Checks on Section II

In Section II, we considered trends in the frequency with which different groups of words were mentioned in our newspaper text. We showed that the share of words related to routine cognitive and routine manual tasks have declined over the sample period, while words related to nonroutine tasks have increased in frequency. Moreover, nearly all of these changes have occurred within, rather than between, occupations. In this appendix, we assess the sensitivity of these results to applying different ways weighting occupations, to applying different normalizations for the task measures, to excluding words from our CBOW algorithm, and to examining different subsamples. We summarize the results of these robustness checks in Table 12. This table indicates that, across a wide variety of specifications, within-occupation shifts account for a majority of the overall changes in the types of tasks that workers perform. In the remainder of this appendix, we present the full decomposition results from each of the exercises summarized in Table 12.

In Section II, our decompositions applied either a 6-digit occupation classification or a job-title based classification. In Table 13, we re-compute our decompositions using 4-digit SOC codes as our occupational category. We do so motivated by the relatively coarse occupational categorization at which certain versions and vintages of publicly available data sets categorize occupations.[53] Table 13 indicates that our decomposition results are similar when using either 4-digit or 6-digit SOCs as the occupation classification.

[53]For instance, versions of the American Community Survey—particularly so for vintages from the early 2000s—group multiple different 6-digit SOCs with one another.

Table 13: Trends in Keyword Frequencies: 4-Digit SOC

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1950 Level	2.76				5.11			
	(0.03)				(0.07)			
1950-2000	3.20	2.76	0.44	0.86	2.38	1.96	0.42	0.82
	(0.07)	(0.11)	(0.07)	(0.02)	(0.11)	(0.15)	(0.06)	(0.03)
	C. Nonroutine Manual				D. Routine Cognitive			
1950 Level	0.90				1.85			
	(0.03)				(0.03)			
1950-2000	-0.15	-0.13	-0.02	0.89	-1.02	-1.06	0.04	1.04
	(0.04)	(0.04)	(0.02)	(0.16)	(0.03)	(0.04)	(0.02)	(0.02)
	E. Routine Manual							
1950 Level	0.99							
	(0.03)							
1950-2000	-0.93	-0.89	-0.04	0.96				
	(0.03)	(0.03)	(0.01)	(0.01)				

Notes: See the notes for Table 3. Here, we use 4-digit SOC codes as opposed to 6-digit SOC codes as the occupational unit.

In our benchmark decompositions, we have used data from the decennial census to construct employment weights: Across 4-digit occupations, the weight of each occupation corresponds to the share of full-time workers employed in the occupation. In Table 14, we instead use the share of vacancies in job title j as our measure for ϑ_{jt}. As with our benchmark calculations in Table 4, the largest task shifts occurred away from routine manual tasks and toward nonroutine analytic tasks. The "Within Share" is now smaller for nonroutine analytic, nonroutine interactive, and routine cognitive tasks, and somewhat larger for the other two task measures. Overall, summing across the five measures, approximately 79 percent of the overall shifts in tasks have occurred within job titles.

Throughout the paper, we have pooled ads from our different samples of newspapers. Potentially, however, the trends in task mentions among the ads in the *Boston Globe* may differ from those in the *New York Times* or *Wall Street Journal.* Likewise, trends in task mentions among display ads and classify ads may differ from one another. In Tables 15 and 16, we perform our decomposition for two of the five subsamples of ads. (These two subsamples were present throughout our sample.) Comparing Tables 15 and 16, display ads contain a greater frequency of words referring to nonroutine analytic, nonroutine interactive, and routine manual task words, and a lesser frequency of words referring to other tasks. However, among both subsamples and similar to the pooled sample, the largest shifts occurred away from routine manual tasks and toward nonroutine analytic tasks. Also similar across the three subsamples, within job-title shifts in task mentions account for the primary share of overall task shifts. The average "Within Share" is 89 percent in Table 15 and 90 percent in Table 16.

When we produced Table 3, Table 4, and our other decomposition tables, we applied

Table 14: Trends in Keyword Frequencies: Vacancy Weights

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1950 Level	3.46				5.58			
	(0.03)				(0.11)			
1950-2000	2.82	1.80	1.01	0.64	2.26	1.07	1.19	0.47
	(0.06)	(0.14)	(0.13)	(0.05)	(0.15)	(0.23)	(0.24)	(0.10)
	C. Nonroutine Manual				D. Routine Cognitive			
1950 Level	0.66				2.64			
	(0.02)				(0.05)			
1950-2000	-0.09	-0.10	0.01	1.15	-1.68	-0.95	-0.73	0.56
	(0.02)	(0.04)	(0.04)	(0.53)	(0.05)	(0.15)	(0.13)	(0.08)
	E. Routine Manual							
1950 Level	0.64							
	(0.02)							
1950-2000	-0.60	-0.55	-0.05	0.92				
	(0.02)	(0.02)	(0.02)	(0.03)				

Notes: See the notes for Table 3. In Table 3, we weight 4-digit SOC codes according to their employment. In this table, instead, we do not apply this weighting method. Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

Table 15: Trends in Keyword Frequencies: *New York Times* Classified Ads

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1950 Level	2.85				4.99			
	(0.03)				(0.07)			
1950-2000	2.56	2.16	0.40	0.84	2.08	1.79	0.29	0.86
	(0.06)	(0.28)	(0.27)	(0.10)	(0.12)	(0.29)	(0.29)	(0.14)
	C. Nonroutine Manual				D. Routine Cognitive			
1950 Level	0.98				2.00			
	(0.03)				(0.03)			
1950-2000	-0.23	-0.20	-0.02	0.90	-1.10	-0.95	-0.15	0.86
	(0.04)	(0.06)	(0.05)	(0.24)	(0.03)	(0.08)	(0.07)	(0.06)
	E. Routine Manual							
1950 Level	0.91							
	(0.04)							
1950-2000	-0.85	-0.77	-0.08	0.91				
	(0.04)	(0.04)	(0.02)	(0.03)				

Notes: See the notes for Table 3. Here, we compute our task measure only from the set of classified ads published in the *New York Times.* Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

Table 16: Trends in Keyword Frequencies: *New York Times* Display Ads

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1950 Level	4.27				7.43			
	(0.41)				(1.05)			
1950-2000	2.87	0.43	2.44	0.15	2.00	-1.14	3.13	-0.57
	(0.55)	(0.83)	(0.74)	(0.26)	(1.29)	(1.34)	(1.47)	(57.66)
	C. Nonroutine Manual				D. Routine Cognitive			
1950 Level	0.63				0.84			
	(0.15)				(0.23)			
1950-2000	0.56	0.66	-0.10	1.18	-0.37	-0.45	0.07	1.19
	(0.20)	(0.77)	(0.69)	(2.19)	(0.24)	(0.34)	(0.19)	(1.41)
	E. Routine Manual							
1950 Level	1.88							
	(0.53)							
1950-2000	-1.84	-1.83	-0.00	1.00				
	(0.53)	(0.53)	(0.05)	(0.03)				

Notes: See the notes for Table 3. Here, we compute our task measure only from the set of display ads published in the *New York Times*. Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

not only the 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 17, we recompute trends in keyword frequencies, now using Spitz-Oener (2006)'s original mapping between words and task groups. In this table, trends in keyword frequencies—increasing for nonroutine analytic and interactive tasks, decreasing for routine tasks, and little change for nonroutine manual tasks—are similar to those depicted in Table 4. Moreover, as in Table 4, a large fraction of the overall changes in keyword frequencies occur within occupations.

Next, a possible limitation is that we are using job ads (which capture the task content of newly formed jobs) to measure stock of jobs existing at that point in time. Underlying our results is the assumption that job ads reflect the task content of all jobs within the occupation (both new and existing). Here, we make the opposite assumption, namely that once a job is formed its task content is fixed over time. With this assumption, and with knowledge of the rate at which jobs turn over within an occupation (call this ς_j), we can compute the evolution of the task content of the stock ($\check{T}_{jt}$) of a given task in occupation j using a perpetual inventory method:

$$\check{T}_{jt} = \widetilde{T}_{jt} \cdot \varsigma_j + \check{T}_{jt-1} \cdot (1 - \varsigma_j), \tag{6}$$

initializing $\check{T}_{jt} = \widetilde{T}_{jt}$ for $t = 1950$. In this equation $\widetilde{T}_{jt}$, equals a measure of the task content of occupation j in job ads at time t. To measure the job turnover rate, we take the sample (within the CPS-ASEC) of employed workers.[54] For workers employed in occupation j, we compute

[54]We restrict attention to workers who are between the age of 16 and 65, who work at least 40 weeks in the previous

Table 17: Trends in Keyword Frequencies, No CBOW

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1950 Level	1.65				1.76			
	(0.03)				(0.04)			
1950-2000	0.95	0.71	0.24	0.75	0.37	0.16	0.21	0.43
	(0.04)	(0.12)	(0.11)	(0.12)	(0.05)	(0.10)	(0.10)	(0.28)
	C. Nonroutine Manual				D. Routine Cognitive			
1950 Level	0.27				0.38			
	(0.01)				(0.01)			
1950-2000	-0.01	-0.05	0.04	6.11	-0.30	-0.29	-0.01	0.95
	(0.02)	(0.03)	(0.03)	(10.37)	(0.01)	(0.02)	(0.02)	(0.05)
	E. Routine Manual							
1950 Level	0.26							
	(0.01)							
1950-2000	-0.25	-0.24	-0.00	0.98				
	(0.01)	(0.01)	(0.01)	(0.03)				

Notes: See the notes for Table 3. In Table 3, 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. Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

the turnover rate as the fraction of workers who were either in a different occupation in the previous year or who had more than one employer. Once we have computed $\check{T}_{jt}$, we recompute the overall and within-occupation changes in tasks in Table 18. The results from this table are that (i) the overall shifts in tasks are smaller than in the benchmark specification, and (ii) the "Within Share" is slightly smaller than in the benchmark specification, yet still above 70 percent for each task measure.

Finally, throughout the paper, our task measures were stated as task mentions divided per 1000 job ad words. As we have discussed in Section II.B, our benchmark measures are not directly comparable with one another. Using T_{jt}^h to refer to the average number of mentions of task h in ads for job title j in year t, as an alternative to our benchmark measures, we place our task measures on a comparable scale via the following equations:

$$\widetilde{T}_{jt}^h = \frac{\widehat{T}_{jt}^h}{\sum_{h' \in \text{Spitz-Oener Tasks}} \widehat{T}_{jt}^{h'}} \text{, where} \tag{7}$$

$$\widehat{T}_{jt}^h = \frac{T_{jt}^h}{\sum_{t'=1950}^{2000} \sum_{j' \in \text{JobTitles}} S_{j't'} T_{j't'}^h} \text{, and} \tag{8}$$

S_{jt} equals the fraction of year t ads which have j as the job title. Within Equation 8, our first transformation requires that the Spitz-Oener (2006) tasks have the same mean. Our second

year, who have non-allocated information on age, race, gender, occupation, and number employers.

Table 18: Trends in Keyword Frequencies: Perpetual Inventory-based Weighting

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1950 Level	2.86				5.03			
	(0.03)				(0.07)			
1950-2000	2.51	1.87	0.64	0.74	2.01	1.49	0.51	0.74
	(0.06)	(0.23)	(0.21)	(0.08)	(0.12)	(0.25)	(0.25)	(0.12)
	C. Nonroutine Manual				D. Routine Cognitive			
1950 Level	0.97				1.99			
	(0.03)				(0.03)			
1950-2000	-0.21	-0.18	-0.03	0.85	-1.10	-0.92	-0.18	0.83
	(0.04)	(0.06)	(0.05)	(0.24)	(0.03)	(0.06)	(0.06)	(0.05)
	E. Routine Manual							
1950 Level	0.91							
	(0.03)							
1950-2000	-0.84	-0.75	-0.09	0.89				
	(0.04)	(0.04)	(0.02)	(0.02)				

Notes: See the notes for Table 3. In comparison, we here apply Equation 6 to impute occupations' task measures. Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

transformation, in Equation 7, then ensures that the task measures for a given job title sum to 1. (A disadvantage of this normalization is that $\widetilde{T}^{h}_{jt}$ will be ill-defined for job title-year combinations for which $\sum_{h' \in \text{Spitz-Oener Tasks}} \widehat{T}^{h'}_{jt}$=0. This occurs for certain job titles which appear only in a few ads in the year t.) In our final set of robustness checks, we assess the sensitivity of our decompositions' results to these normalization. According to the first panel of Table 19, mentions of the nonroutine analytic task share more than doubled between 1950 and 2000, from 15 percent to 33 percent. The nonroutine interactive words also increased, but to a lesser degree. Conversely, the routine manual tasks share substantially declined, decreasing from 19 percent to 2 percent. The decline of routine cognitive tasks is also considerable, going from 27 percent to 14 percent. As with 4, these shifts have primarily occurred within rather than between job titles. Averaging over the five task groups, 76 percent of the overall shift in task mentions have occurred within job titles. (This is somewhat lower than the 88 percent figure that would correspond to the Table 4 decompositions.) Likewise, Table 20 indicates that large "Within Shares" obtain whether one applies our Equation 7-8 normalization to our Deming and Kahn (2018) measures or not. One difference as a result of the normalization: Some of Deming and Kahn (2018) measures decline over our sample period when applying Equations 7 and 8. In contrast, in Table 5 the frequency of mentions of all words related to Deming and Kahn (2018)'s measure increased between 1950 and 2000.

Table 19: Trends in Keyword Shares

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Nonroutine Analytic				B. Nonroutine Interactive			
1950 Level	0.148				0.233			
	(0.003)				(0.003)			
1950-1960	0.051	-0.001	0.052	-0.029	0.038	-0.025	0.063	-0.669
	(0.004)	(0.004)	(0.003)	(0.076)	(0.004)	(0.004)	(0.003)	(0.207)
1960-1970	0.034	0.027	0.007	0.801	0.015	0.006	0.009	0.413
	(0.004)	(0.004)	(0.003)	(0.075)	(0.003)	(0.004)	(0.004)	(0.226)
1970-1980	0.014	0.003	0.011	0.220	0.020	0.015	0.004	0.773
	(0.004)	(0.006)	(0.006)	(0.531)	(0.004)	(0.009)	(0.007)	(0.344)
1980-1990	0.037	0.040	-0.003	1.086	0.061	0.064	-0.004	1.058
	(0.005)	(0.011)	(0.008)	(0.231)	(0.005)	(0.011)	(0.010)	(0.165)
1990-2000	0.051	0.080	-0.029	1.564	-0.003	0.003	-0.007	-0.909
	(0.006)	(0.014)	(0.014)	(0.293)	(0.004)	(0.015)	(0.015)	(7.819)
1950-2000	0.187	0.149	0.038	0.797	0.130	0.064	0.066	0.490
	(0.004)	(0.014)	(0.013)	(0.071)	(0.005)	(0.011)	(0.011)	(0.084)
	C. Nonroutine Manual				D. Routine Cognitive			
1950 Level	0.159				0.267			
	(0.004)				(0.004)			
1950-1960	0.001	0.016	-0.015	17.791	-0.063	-0.010	-0.053	0.155
	(0.007)	(0.007)	(0.002)	(76.382)	(0.006)	(0.007)	(0.003)	(0.099)
1960-1970	0.011	0.023	-0.012	2.093	-0.012	-0.004	-0.007	0.366
	(0.005)	(0.007)	(0.005)	(4.277)	(0.005)	(0.007)	(0.005)	(0.542)
1970-1980	0.046	0.057	-0.011	1.229	-0.046	-0.048	0.002	1.045
	(0.004)	(0.011)	(0.010)	(0.233)	(0.004)	(0.008)	(0.007)	(0.164)
1980-1990	-0.066	-0.082	0.015	1.234	0.019	0.059	-0.040	3.160
	(0.004)	(0.014)	(0.012)	(0.184)	(0.006)	(0.010)	(0.008)	(4.775)
1990-2000	-0.008	0.004	-0.013	-0.530	-0.020	-0.063	0.043	3.109
	(0.005)	(0.017)	(0.016)	(6.796)	(0.007)	(0.019)	(0.017)	(1.424)
1950-2000	-0.016	0.019	-0.035	-1.134	-0.122	-0.067	-0.056	0.545
	(0.005)	(0.013)	(0.011)	(0.901)	(0.004)	(0.017)	(0.015)	(0.132)
	E. Routine Manual							
1950 Level	0.193							
	(0.005)							
1950-1960	-0.027	0.020	-0.047	-0.763				
	(0.007)	(0.008)	(0.003)	(0.729)				
1960-1970	-0.049	-0.052	0.004	1.072				
	(0.004)	(0.007)	(0.004)	(0.094)				
1970-1980	-0.035	-0.028	-0.007	0.798				
	(0.004)	(0.007)	(0.006)	(0.169)				
1980-1990	-0.049	-0.081	0.031	1.632				
	(0.003)	(0.006)	(0.005)	(0.106)				
1990-2000	-0.019	-0.024	0.005	1.266				
	(0.003)	(0.006)	(0.005)	(0.258)				
1950-2000	-0.178	-0.165	-0.014	0.923				
	(0.005)	(0.006)	(0.003)	(0.016)				

Notes: See the notes for Table 3. In this table, we apply the normalizations given in Equations 7 and 8. Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

Table 20: Trends in Keyword Shares: Deming and Kahn (2018) Measures

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	A. Character				B. Computer			
1950 Level	0.202				0.027			
	(0.003)				(0.001)			
1950-1960	-0.027	-0.036	0.009	1.323	0.030	0.022	0.008	0.725
	(0.004)	(0.005)	(0.002)	(0.133)	(0.002)	(0.002)	(0.001)	(0.045)
1960-1970	0.027	0.020	0.007	0.745	0.016	0.023	-0.007	1.460
	(0.003)	(0.004)	(0.003)	(0.114)	(0.001)	(0.003)	(0.002)	(0.158)
1970-1980	0.016	0.033	-0.017	2.034	0.022	0.026	-0.004	1.165
	(0.004)	(0.006)	(0.005)	(0.360)	(0.001)	(0.003)	(0.003)	(0.156)
1980-1990	-0.051	-0.056	0.006	1.116	0.031	0.050	-0.019	1.605
	(0.004)	(0.010)	(0.009)	(0.168)	(0.002)	(0.007)	(0.007)	(0.249)
1990-2000	-0.045	-0.058	0.013	1.298	0.041	0.017	0.024	0.414
	(0.003)	(0.008)	(0.008)	(0.168)	(0.003)	(0.009)	(0.009)	(0.209)
1950-2000	-0.079	-0.097	0.018	1.226	0.140	0.137	0.003	0.980
	(0.004)	(0.006)	(0.005)	(0.065)	(0.002)	(0.005)	(0.005)	(0.034)
	C. Customer Service				D. Financial			
1950 Level	0.149				0.149			
	(0.002)				(0.002)			
1950-1960	-0.006	-0.017	0.011	2.782	-0.022	-0.016	-0.006	0.717
	(0.002)	(0.002)	(0.001)	(1.719)	(0.003)	(0.004)	(0.001)	(0.086)
1960-1970	-0.011	-0.012	0.001	1.130	-0.010	-0.014	0.005	1.488
	(0.002)	(0.002)	(0.002)	(0.218)	(0.003)	(0.003)	(0.002)	(0.316)
1970-1980	0.004	0.009	-0.006	2.688	-0.033	-0.043	0.010	1.300
	(0.002)	(0.003)	(0.003)	(3.135)	(0.002)	(0.003)	(0.002)	(0.062)
1980-1990	0.015	0.010	0.004	0.707	0.001	-0.004	0.005	-5.623
	(0.003)	(0.004)	(0.004)	(0.251)	(0.002)	(0.003)	(0.002)	(803.716)
1990-2000	-0.008	-0.010	0.002	1.303	0.002	0.001	0.001	0.613
	(0.002)	(0.006)	(0.006)	(0.913)	(0.002)	(0.004)	(0.004)	(7.175)
1950-2000	-0.006	-0.019	0.013	3.074	-0.062	-0.076	0.014	1.229
	(0.003)	(0.006)	(0.005)	(14.792)	(0.003)	(0.004)	(0.004)	(0.062)
	E. People Management				F. Problem Solving			
1950 Level	0.112				0.102			
	(0.001)				(0.002)			
1950-1960	0.014	0.014	0.000	0.975	0.019	0.025	-0.006	1.314
	(0.002)	(0.003)	(0.002)	(0.125)	(0.003)	(0.004)	(0.002)	(0.101)
1960-1970	0.037	0.042	-0.005	1.149	-0.025	-0.019	-0.006	0.762
	(0.003)	(0.004)	(0.003)	(0.072)	(0.002)	(0.003)	(0.003)	(0.115)
1970-1980	-0.019	-0.038	0.018	1.955	-0.015	-0.019	0.004	1.291
	(0.002)	(0.003)	(0.002)	(0.157)	(0.002)	(0.004)	(0.003)	(0.215)
1980-1990	-0.030	-0.025	-0.006	0.816	-0.005	-0.011	0.006	2.185
	(0.002)	(0.008)	(0.007)	(0.257)	(0.002)	(0.004)	(0.004)	(1.753)
1990-2000	-0.028	-0.024	-0.004	0.855	0.001	-0.004	0.005	-2.539
	(0.002)	(0.009)	(0.009)	(0.304)	(0.002)	(0.005)	(0.005)	(23.262)
1950-2000	-0.027	-0.031	0.003	1.128	-0.025	-0.028	0.004	1.152
	(0.002)	(0.004)	(0.004)	(0.147)	(0.002)	(0.004)	(0.003)	(0.131)

Notes: Continued on the following page.

Table 20 (Continued): Trends in Keyword Shares: Deming and Kahn (2018) Task Measures

	Total	Within	Between	Within Share	Total	Within	Between	Within Share
	G. Project Management				H. Social			
1950 Level	0.111 (0.002)				0.045 (0.001)			
1950-1960	0.007 (0.002)	-0.009 (0.002)	0.016 (0.001)	-1.366 (1.501)	-0.003 (0.002)	0.001 (0.002)	-0.004 (0.001)	-0.200 (5.626)
1960-1970	-0.009 (0.002)	-0.005 (0.002)	-0.004 (0.002)	0.568 (0.208)	0.005 (0.001)	0.008 (0.003)	-0.004 (0.002)	1.825 (0.491)
1970-1980	0.001 (0.002)	-0.004 (0.003)	0.005 (0.003)	-2.775 (47.545)	0.022 (0.001)	0.020 (0.004)	0.002 (0.003)	0.917 (0.147)
1980-1990	-0.016 (0.001)	-0.011 (0.004)	-0.004 (0.003)	0.712 (0.235)	0.038 (0.001)	0.044 (0.004)	-0.006 (0.004)	1.157 (0.116)
1990-2000	0.012 (0.002)	0.043 (0.014)	-0.031 (0.014)	3.507 (1.127)	0.019 (0.003)	0.016 (0.008)	0.002 (0.007)	0.880 (0.373)
1950-2000	-0.005 (0.002)	0.014 (0.013)	-0.019 (0.013)	-3.065 (17.250)	0.080 (0.003)	0.090 (0.008)	-0.010 (0.006)	1.121 (0.076)
	I. Writing							
1950 Level	0.102 (0.002)							
1950-1960	-0.012 (0.003)	0.015 (0.003)	-0.027 (0.002)	-1.265 (0.615)				
1960-1970	-0.029 (0.002)	-0.042 (0.004)	0.013 (0.003)	1.466 (0.087)				
1970-1980	0.002 (0.002)	0.015 (0.006)	-0.013 (0.006)	7.759 (23.697)				
1980-1990	0.017 (0.002)	0.004 (0.007)	0.013 (0.007)	0.223 (0.432)				
1990-2000	0.006 (0.003)	0.019 (0.007)	-0.013 (0.007)	3.194 (18.920)				
1950-2000	-0.016 (0.003)	0.010 (0.006)	-0.027 (0.005)	-0.651 (0.446)				

Notes: See the notes for Table 3. In this table, we apply the normalizations given in Equations 7 and 8. Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

F Robustness Checks on Section III

This appendix compiles two robustness checks relative to our analysis in Section III. First, we reproduce Table 6 using the normalizations introduced in Equations 7 and 8. Second, we re-estimate the relationship between job title vintages and tasks while controlling for year fixed effects.

Table 6 in the body of the paper characterized the evolution of Manager, Machinist, Cashier, and Real Estate Sales jobs by comparing each job title's task frequencies to the task frequencies in other job titles in our data set. In these calculations, we applied our baseline task measures, which is the frequency of Spitz-Oener (2006) task mentions per 1000 job ad words. Here, we assess the sensitivity of Table 6 to this measure of task intensity. Alternatively, Table applies Equations 7 and 8 to construct measures of the shares of each each Spitz-Oener (2006) task among the set of ads for each job title - decade pair.

Using these normalizations, we compute that 1950s Manager jobs were 12 percent nonroutine analytic, 20 percent nonroutine interactive, 17 percent nonroutine manual, 15 percent routine cognitive, and 36 percent routine manual. Managerial jobs in the 1950s closely mirrored—according to this five-dimensional representation—Pressman jobs. Averaging over all ads in our sample period, ads with Pressman as the job title were 14 percent nonroutine analytic, 19 percent nonroutine interactive, 20 percent nonroutine manual, 12 percent routine cognitive, and 36 percent routine manual. Over time, the nonroutine analytic and interactive task content of "Manager" jobs increased. Correspondingly, we find that 1960s Manager job ads were similar to Purchasing Agent ads. Later in the sample, Managers more closely resembled Editors (in the 1980s) and Recruiters (in the 1990s). The remaining three panels of Table 6 characterize the evolution of Machinist, Cashier, and Real Estate Sales job ads. Machinist job ads only meaningfully shifted between the 1980s and 1990s. In the last decade of our sample, Machinist ads began to resemble 1950-2000 Printer job ads. The Cashier job ads from the 1950s contained similar task combinations to Comptometer Operator ads. By the 1990s, Cashier job ads more closely resembled ads for Secretary Receptionists.

These patterns qualitatively mirror those in Table 6. Managers more closely mimic production-related occupations (i.e., Production Manager or Pressman) in the 1950s and jobs centered around interpersonal tasks (Recruiter or Coordinator) in the 1990s. Machinists are relatively unchanged in their task mix throughout the first four decades of our sample. There are also noteworthy differences: According to Table 6, Real Estate Sales only noticeably changed in the 1990s, when this job title's task mix approximated that of 1950-2000 Furniture Salespersons. In the bottom panel of Table 21 instead, Real Estate Sales closely mirrored Furniture Salespersons in the 1950s, Telephone Sales in the 1960s and 1970s, and Advertising Sales in the 1980s.

In Section III.B, we document that the task content of job titles that emerge later in our sample period differs from the content of those that emerge earlier. Our key finding is that newer job titles have higher intensities of nonroutine analytic and interactive tasks, lower intensities of routine cognitive and manual tasks, and higher intensities of computer skills.

Table 21: Near Job Titles

	Shares	Similar Job Title	Shares of Similar Job Title
Panel A: Manager			
1950-1959	(0.12, 0.20, 0.17, 0.15, 0.36)	Pressman	(0.14, 0.19, 0.20, 0.12, 0.36)
1960-1969	(0.21, 0.24, 0.19, 0.12, 0.24)	Purchasing Agent	(0.17, 0.29, 0.17, 0.09, 0.27)
1970-1979	(0.29, 0.29, 0.19, 0.09, 0.15)	Manager	(0.27, 0.30, 0.18, 0.12, 0.13)
1980-1989	(0.29, 0.33, 0.19, 0.11, 0.08)	Editor	(0.30, 0.32, 0.18, 0.11, 0.10)
1990-2000	(0.32, 0.36, 0.18, 0.12, 0.02)	Recruiter	(0.32, 0.35, 0.17, 0.13, 0.03)
Panel B: Machinist			
1950-1959	(0.03, 0.03, 0.18, 0.05, 0.71)	Machinist	(0.03, 0.02, 0.15, 0.04, 0.75)
1960-1969	(0.03, 0.02, 0.14, 0.05, 0.76)	Machinist	(0.03, 0.02, 0.15, 0.04, 0.75)
1970-1979	(0.03, 0.01, 0.11, 0.01, 0.84)	Machinist	(0.03, 0.02, 0.15, 0.04, 0.75)
1980-1989	(0.04, 0.02, 0.16, 0.03, 0.76)	Machinist	(0.03, 0.02, 0.15, 0.04, 0.75)
1990-2000	(0.14, 0.12, 0.25, 0.03, 0.45)	Printer	(0.12, 0.15, 0.18, 0.10, 0.46)
Panel C: Cashier			
1950-1959	(0.06, 0.08, 0.07, 0.64, 0.15)	Comptometer Operator	(0.07, 0.10, 0.06, 0.62, 0.15)
1960-1969	(0.08, 0.09, 0.11, 0.55, 0.16)	Clerk	(0.07, 0.10, 0.08, 0.58, 0.17)
1970-1979	(0.07, 0.09, 0.24, 0.43, 0.18)	Teller	(0.14, 0.14, 0.16, 0.44, 0.13)
1980-1989	(0.11, 0.20, 0.17, 0.46, 0.07)	Claims	(0.16, 0.17, 0.13, 0.44, 0.10)
1990-2000	(0.21, 0.30, 0.06, 0.38, 0.05)	Secretary Receptionist	(0.15, 0.26, 0.08, 0.44, 0.06)
Panel D: Real Estate Sales			
1950-1959	(0.10, 0.63, 0.13, 0.03, 0.11)	Furniture Salesperson	(0.14, 0.54, 0.14, 0.07, 0.11)
1960-1969	(0.09, 0.69, 0.11, 0.07, 0.04)	Telephone Sales	(0.08, 0.70, 0.11, 0.07, 0.05)
1970-1979	(0.07, 0.74, 0.12, 0.02, 0.05)	Telephone Sales	(0.08, 0.70, 0.11, 0.07, 0.05)
1980-1989	(0.08, 0.74, 0.11, 0.06, 0.01)	Advertising Sales	(0.08, 0.76, 0.10, 0.04, 0.02)
1990-2000	(0.09, 0.61, 0.23, 0.07, 0.00)	Real Estate Sales	(0.09, 0.69, 0.14, 0.05, 0.04)

Notes: See the notes related to Table 6. In contrast to Table 6, we apply the normalizations given in Equations 7 and 8. Further, we use job titles as opposed to 6-digit SOC codes as the occupational unit.

These patterns hold both within and across conventionally defined SOC occupation codes.

To estimate Equation 2, we calculate task contents for each job title over our entire sample period, and hence the time time periods we use vary across job titles. The Section III.B specification is our preferred one. Much of the overall time trends in tasks is due to the arrival of new vintages of job titles. We prefer our estimates to reflect that channel.

However, one may wish to compare newer and older vintages at the same point in time, since other time trends in tasks may account for part of the observed differences between newer and older job titles. In this section we re-estimate Equation 2 at the job title-year level, and include year fixed effects. The comparison below is between newer and older job titles, observed in the same year.

We estimate the following regression:

$$\widetilde{T}_{jt}^{h} = \beta_o + \beta_t + \beta_1 v_j^p + \varepsilon_{jht} \tag{9}$$

Unlike Equation 2, Equation 9 exploits variation across job titles and over time. In this

Table 22: Relationship Between Task Measures and Job Title Vintages

Dependent Variable	—— Nonroutine —— Analytic	Interactive	Manual	— Routine — Cognitive	Manual	Deming and Kahn Computer
Panel A: No SOC Fixed Effects, $p = 0.05$						
Coefficient	0.007	0.013	0.001	-0.015	-0.001	0.023
Standard Error	0.002	0.002	0.000	0.001	0.000	0.001
Panel B: 6-Digit SOC Fixed Effects, $p = 0.05$						
Coefficient	-0.011	0.003	-0.001	-0.007	-0.002	0.005
Standard Error	0.001	0.001	0.000	0.000	0.000	0.001
Panel C: No Fixed Effects, $p = 0.50$						
Coefficient	0.020	0.056	0.002	-0.027	-0.004	0.017
Standard Error	0.002	0.003	0.000	0.002	0.000	0.002
Panel D: 6-Digit SOC Fixed Effects, $p = 0.50$						
Coefficient	0.006	0.024	0.001	-0.005	-0.001	0.007
Standard Error	0.001	0.002	0.000	0.001	0.000	0.001
Panel E: No Fixed Effects, $p = 0.95$						
Coefficient	0.002	0.013	-0.001	0.007	-0.002	-0.002
Standard Error	0.002	0.002	0.000	0.002	0.000	0.001
Panel F: 6-Digit SOC Fixed Effects, $p = 0.95$						
Coefficient	0.008	0.008	0.001	0.008	0.001	0.003
Standard Error	0.001	0.001	0.000	0.001	0.000	0.001

Notes: Within each panel and column, we present coefficient estimates and standard errors corresponding to estimates of Equation 22. In this table, each observation is a job-title year combination. This contrasts with Table 7, in which each job title corresponds to a single observation.

regression, we include both occupation fixed effects (β_o) and year fixed effects (β_t). For most task and job title vintage measures, the results in Table 22 agree with those of Table 7. Nonroutine analytic tasks, nonroutine interactive tasks, and computer skills are more frequently mentioned in newer job titles. Routine manual tasks are more frequently mentioned in older vintage job titles. Not surprisingly, across all specifications, the estimated coefficients are somewhat smaller in absolute value than those in Table 7. Moreover, we note that the task versus vintage relationship is at times sensitive to the measure of job title vintage under consideration.

Appendix References

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.

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

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.

Deming, D., and L. B. Kahn (2018): "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals," *Journal of Labor Economics*, 36(1), 337–369.

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.

Hansen, S., M. McMahon, and A. Prat (2018): "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach," *Quarterly Journal of Economics*, 133(2), 801–870.

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.

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.

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

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.