

## Rising Top, Falling Bottom: Industries and Rising Wage Inequality<sup>†</sup>

By JOHN HALTIWANGER, HENRY R. HYATT, AND JAMES R. SPLETZER\*

*Most of the rise in overall earnings inequality from 1996 to 2018 is accounted for by rising between-industry dispersion. The contribution of industries is right-skewed with the top 10 percent of four-digit NAICS industries dominating. The top 10 percent are clustered in high-paying high-tech and low-paying retail sectors. In the top industries, high-wage workers are increasingly sorted to high-wage industries with rising industry premia. In the bottom industries, low-wage workers are increasingly sorted into low-wage industries, with rising employment and falling industry wage premia. (JEL J23, J24, J31, L25, M52)*

A growing number of studies attribute increases in earnings inequality to rising between-firm dispersion.<sup>1</sup> We confirm this pattern with US matched employer-employee data from 1996 to 2018. Our contribution is to demonstrate that rising between-industry dispersion accounts for most of the overall increase in earnings inequality, and is driven by a relatively small number of industries in the tails of the industry earnings distribution. About 10 percent of four-digit NAICS industries account for virtually all of the increase in between-industry dispersion, while

\*Haltiwanger: University of Maryland (email: halt@umd.edu); Hyatt: Center for Economic Studies, US Census Bureau (email: henry.r.hyatt@gmail.com); Spletzer: Retired (email: james.r.spletzer@gmail.com). Emi Nakamura was the coeditor for this article. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the US Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed (CBDRB-FY21-046, CBDRB-FY21-143, CBDRB-FY21-155, CBDRB-FY21-323, CBDRB-FY22-356, CBDRB-FY23-0396, and CBDRB-FY23-0399). We thank Keith Bailey, Nicholas Bloom, Andrew Foote, and participants at the 2021 Wage Dynamics in the 21st Century conference, the 2021 Allied Social Science Associations meeting, the 2022 Washington Area Labor Economics Symposium, the 2022 Research Data Center conference, the 2022 Society of Labor Economists conference, the 2022 Society for Economic Dynamics conference, the 2022 NBER Summer Institute, and the 2022 Federal Reserve Bank of Atlanta Center for Human Capital Studies conference for helpful comments, two anonymous referees, and we especially thank our discussants Salomé Baslandze, Christina Patterson, and Ayşegül Şahin. We also thank Matilde Serrano for excellent research assistance. John Haltiwanger was also a Schedule A employee of the US Census Bureau at the time of the writing of this paper.

<sup>†</sup>Go to <https://doi.org/10.1257/aer.20221574> to visit the article page for additional materials and author disclosure statement(s).

<sup>1</sup>Barth et al. (2016) and Song et al. (2019) provide evidence for the United States. These papers follow an earlier literature emphasizing the importance of rising between-firm effects for earnings inequality that includes Davis and Haltiwanger (1991) and Dunne et al. (2004). Card, Heining, and Kline (2013) and Card, Cardoso, and Kline (2016) consider the role of firms in rising inequality in Germany and Portugal, respectively. We note that in both the title of the paper and throughout the paper, we often use the term *wage* as a shorthand for earnings per worker. All of our analysis is about earnings per worker on an annual basis consistent with the recent literature using administrative data. Our core administrative data infrastructure does not include information on hours per worker.

accounting for less than 40 percent of employment.<sup>2</sup> Remarkably, the remaining 90 percent of four-digit industries contribute little to rising between-industry earnings inequality.<sup>3</sup>

The top 10 percent of industries that contribute to rising inequality include 19 that are high-paying (top quartile in pay). These industries account for 54.1 percent of the increase in between-industry inequality. Average worker earnings have surged in these industries accompanied by employment gains. The top 3 of these are high-paying, high-tech industries: Software Publishers (NAICS 5112), Computer Systems Design (5415), and Other Information Services (5191)—and, in total, 11 of these 19 high-paying industries are high-tech. As discussed in Oliner, Sichel, and Stiroh (2007) and Fernald (2014), these industries are characterized as the source of rapid technological advances over our sample period.

Eleven low-paying (bottom quartile) industries constitute the remainder of the top 10 percent in terms of their contribution to inequality. These industries, in combination, account for 44.1 percent of the increase in between-industry inequality. More than one-fourth of the increase is accounted for by just 3 of these 11: Restaurants and Other Eating Places (7225), Other General Merchandise Stores (4529), and Grocery Stores (4451). These industries have gone through substantial changes in recent decades, moving away from single-establishment firms to large, national chains, see Foster, Haltiwanger, and Krizan (2006) and Autor et al. (2020). Average earnings declines accompanied by substantial increases in employment in these low-paying industries have led to especially large increases in inequality.<sup>4</sup>

A distinctive feature of the dominant 10 percent of industries is that they exhibit a sharp increase in the share of employment at mega firms, which we define as firms with more than 10,000 employees. Strikingly, the remaining 90 percent of industries exhibit small declines in the share of employment at mega firms. For the low-paying dominant industries, there is a decline in the size-earnings premium. For the high-paying dominant industries, mega firms increased earnings substantially relative to both small firms in the same industry and to earnings of the average industry. Thus, we find that the rise in “superstar” firms (e.g., Autor et al. 2020) is concentrated in these dominant industries with accompanying systematic changes in the size-earnings premia.

The top 30 industries also exhibit distinctive changes in the education and occupations of their workers. Especially notable are large increases in the share

<sup>2</sup>Industry definitions follow the North American Industry Classification System (NAICS). Haltiwanger and Spletzer (2020, 2022) use LEHD data to explore between-industry dispersion. The first paper was a preliminary working paper that has been completely subsumed by the current paper. The second paper considers the impact of changes in industry structure on labor market fluidity. Haltiwanger, Hyatt, and Spletzer (2023) explore distinct but related measurement issues in reconciling differences in earnings inequality from survey and administrative data (see online Appendix D).

<sup>3</sup>There are a small number of industries that are a drag on rising inequality. Partly this reflects manufacturing industries that have high earnings premia and have declining employment. In online Appendix H.2, we present evidence that in earlier eras (1980–1986 to 1996–2002) that the drag from manufacturing was larger but not dominant.

<sup>4</sup>Briskar et al. (2022) replicate several of these key facts using data from Italy during a similar time period. They show that 5 of 524 industries have a between-industry variance contribution greater than 5 percent, while we find 5 of 301. Twenty-six industries have a contribution of greater than 1 percent in the Italian data, while we find 30. In both our datasets, the restaurant industry has the largest between-industry variance contribution: 19.7 percent in the Italian data, and 16.9 percent in ours. The Italian data also shows that high-paying industries in their top 26 are driven by earnings growth, and low-paying industries in their top 26 are driven by employment growth, which is consistent with our findings.

of workers with a bachelor's or advanced degree in the top 19 high-paying industries. These industries also exhibit sharp increases in high-paying occupations such as Business and Financial Operations (SOC 13), Computer and Mathematical Science (15), and strong declines in Office and Administrative Support (43). The top 11 low-paying industries exhibit substantial increases in low-paying occupations such as Food Preparation and Serving Related (35) and Personal Care and Service (39). Strikingly, nearly all of the growth in the employment share of high-paying occupations occurs in the top 19 high-paying industries, and all of the growth in low-paying occupations occurs in the top 11 low-paying industries.

We provide further insights about rising between-industry inequality by quantifying the role of industry premia and between-industry sorting. By sorting, we refer to the frequency with which high (low) wage workers are employed in industries with high (low) wage premia, as well as the concentration of high (low) wage workers in particular industries. We follow Song et al. (2019) in using an Abowd, Kramarz, and Margolis (1999)—hereafter, AKM—decomposition of earnings as our initial method for quantifying these effects. We also characterize these effects using a standard human capital equation (see Hoffman, Lee, and Lemieux 2020) that relates earnings to age, education, occupation, and industry effects. This takes advantage of linked data from the Current Population Survey (CPS). Finally, we describe how earnings vary with industry and occupation using published aggregates from the Occupational Employment and Wage Statistics (OEWS). Regardless of our approach, we find that an increase in sorting plays the most important role. Rising dispersion in between-industry premia also make an important direct contribution in addition to the indirect contribution via sorting.

An increasing covariance between industry premia and earnings premia for worker characteristics for an industry underlies the dominant role of sorting. This increased covariance depends critically on the presence of industry premia which have been an active part of the debate in the literature since at least Krueger and Summers (1988). In this respect, our findings are related to but distinct from both the earlier and more recent literature on inter-industry earnings differentials. Especially related is the recent work of Card, Rothstein, and Yi (2024), who assess the relationship between industry premia and the cross-sectional dispersion in earnings using matched employer-employee data.

Our work is distinct from the inter-industry earnings differential literature in two important and related ways. First, we focus on the change in inequality rather than the level.<sup>5</sup> Second, we distinguish the direct effects of industry premia from the indirect contribution from the covariance of industry premia and the earnings premia for worker characteristics in the industry. Taken together, the direct and indirect effects of industry premia account for most of the dominant role of rising between industry dispersion. Moreover, these effects are concentrated in the top 30 industries.

A rising between-industry covariance requires that industry premia are changing in the same direction as the earnings premia for observable characteristics such as education and occupation, as well as unobservable worker effects. A major finding of our paper is that this property holds using either the AKM approach (for

<sup>5</sup>We find similar cross-sectional industry premia to those found in this literature. For example, Restaurants and Other Eating Places (7225) have very low premia and Software Publishers (5112) have very high premia.

unobservable effects) or education and occupation effects (using the CPS-LEHD and OEWS data). We find, for example, that Other Information Services (5191, where internet portals and search engine companies are located) has increases in both the industry premia and earnings premia for worker effects of over 20 percent over our sample period.

Our findings imply that understanding rising earnings inequality during the last several decades requires understanding the restructuring of how firms organize themselves in a relatively small set of industries. The importance of select industries is well-recognized in the study of productivity but has been neglected in the inequality literature, which has focused on changes in the relative demand and supply of workers with different skills that are used to accomplish different tasks. While our results are not inconsistent with that view, our findings highlight that the skill and task mix varies dramatically across firms and especially industries reflecting the different ways that production is organized. As such, when structural changes such as adoption of new technologies or globalization impact the relative demand of workers by skills and tasks this manifests itself through changes in the organization of production in distinct ways across detailed industries.

Our findings provide a related but distinct perspective on rising polarization (see Autor, Katz, and Kearney 2006, 2008; Goos and Manning 2007; Acemoglu and Autor 2011). We find that the rise in overall earnings inequality from 1996 to 2018 is accounted for by the polarization of industry wage structures in key industries. In these key industries, high-wage workers are increasingly sorted to high-wage industries with rising industry premia and are increasingly working with each other. Low-wage workers are increasingly sorted into low-wage industries, with rising employment and falling industry wage premia and are increasingly working with each other. This role of industry was not detectable in the earlier literature given the limitations of industry codes in the household survey data that has been the workhorse of the inequality literature (see Haltiwanger, Hyatt, and Spletzer 2023). Using high-quality industry codes that are inherent in our administrative matched employer-employee data, we are able to identify and analyze this polarization of industry wage structures.

## I. Data

### A. LEHD Data and the Analysis Sample

For our core analysis, we use Longitudinal Employer-Household Dynamics (LEHD) linked employer-employee data, which is created by the US Census Bureau as part of the Local Employment Dynamics federal-state partnership. The LEHD data are derived from unemployment insurance (UI) wage records and quarterly census of employment and wages (QCEW) data, see Abowd et al. (2009). Every quarter, employers who are subject to state UI laws—approximately 98 percent of all private sector employers, plus state and local governments—are required to report information on their workers (the wage records, which record the quarterly earnings of every worker in the firm) and their workplaces (the QCEW, which provides the industry and location of each establishment). The wage records and the QCEW data submitted by the states to the US Census Bureau are enhanced

with census and survey microdata in order to incorporate information about worker demographics (age and gender) and the firm (firm age and firm size). A job in the LEHD data is defined as a worker-employer match, and earnings is defined as the amount earned from that job during a given quarter.

Because states have joined the LEHD program at different times, the length of the time series varies by state. We use data from 18 states that have data from 1996:I through 2018:IV, which gives us data for 23 years.<sup>6</sup> We create annual person-level data as  $Y_t^i = \sum_j \sum_{q \in t} Y_{qt}^{ij}$ , which sums the earnings  $Y$  that worker  $i$  receives from firm  $j$  in any quarter  $q$  during year  $t$ . We use the federal employer identification number (EIN) as the firm identifier, and include only private sector jobs.<sup>7</sup> We follow Abowd, McKinney, and Zhao (2018) and delete any worker with 12 or more jobs in a given year. A worker's employer is the firm that contributes the most earnings in a given year. This yields 1.395 billion person-year observations (averaging about 61 million persons per year).

We create our analytical dataset following the sample restrictions of Song et al. (2019). We restrict to persons aged 20–60 with annual real (2013 = 100 PCE deflator) earnings greater than \$3,770 (=13 weeks  $\times$  40 hours per week  $\times$  \$7.25 minimum wage). We top-code annual earnings at the 99.999 percent value (for anyone with earnings in the top 0.001 percent, we replace their earnings with the mean earnings of the top 0.001 percent). Our dataset has 1.048 billion person-year observations (an average of 45.6 million persons per year). We use real annual log earnings  $y_t^i = \ln(Y_t^i)$ . We define three seven-year intervals (1996–2002, 2004–2010, 2012–2018), reducing the sample to 959 million person-year observations.<sup>8</sup>

Again following Song et al. (2019), we restrict to firms that employ 20 or more persons in a given year. This reduces our sample to 763 million person-year observations. Due to Census Bureau disclosure rules (to avoid small samples), we further restrict the LEHD data to include firms that have at least one male and one female to allow for separate variance decompositions by gender. The final LEHD data used to create all our results contains 758 million person-year observations.

## B. CPS-LEHD, OEWS, and LBD

We consider additional datasets to supplement our analysis. The first integrates CPS data at the person-level into the LEHD data.<sup>9</sup> The matched CPS-LEHD data provides estimates of changing earnings inequality and the contribution of industries that are very similar to the full LEHD when using LEHD earnings and industry codes. The advantage of this integrated data is that we can exploit observed

<sup>6</sup>These 18 states are California, Colorado, Connecticut, Hawaii, Idaho, Illinois, Kansas, Louisiana, Maryland, Montana, Minnesota, New Jersey, North Carolina, Oregon, Rhode Island, Texas, Washington, and Wisconsin. These 18 states account for roughly 44 percent of national employment. The time series of employment from these 18 states closely tracks the national time series of total private sector employment published by the BLS.

<sup>7</sup>Haltiwanger and Spletzer (2022) estimate variance decompositions using different levels of firm identifiers: the state UI account number, the EIN, and the enterprise. They find that rising between-industry dispersion accounts for most of the rising between-firm inequality regardless of the definition of the firm.

<sup>8</sup>Descriptive statistics are provided in online Appendix A. The analysis in the paper pools males and females. We have conducted separate estimates by gender. Results, which are largely similar for females and males, are available by request.

<sup>9</sup>We provide details of this matching process in online Appendix D.

education and occupation to help understand the role of rising between- industry dispersion to earnings inequality. One potential difference between the full LEHD sample and the CPS-LEHD sample is that we don't make any minimum firm size restrictions. This feature enables us to show that this is not critical as results are robust in removing that restriction.

We also use published OEWS aggregates to explore the role of occupation in rising between-industry dispersion. We obtain a consistent time series for 22 occupation categories for 281 detailed (four-digit) NAICS industries for the years 2002–2016.<sup>10</sup> As we will see, the top industries that contribute to rising dispersion in the OEWS strongly overlap with those in the LEHD data.

We also use establishment-level microdata from the Longitudinal Business Database (LBD) to decompose the variance of real log earnings per worker across firms and industries. We use the LBD to assess the robustness of our LEHD-based analysis for coverage across states, our firm size restriction, and different business definitions (establishment and EIN). We also use the LBD to assess the role of between-industry changes in inequality in earlier periods (1980–1986 to 1996–2002).

### C. Industry Codes

Industry codes play a fundamental role in our analysis. We define industry at the four-digit NAICS level.<sup>11</sup> Our basic results use establishment-level industry codes from the BLS QCEW program. We assign each EIN its dominant industry. If an EIN has  $N > 1$  establishments with  $M$  industry codes, where  $N \geq M > 1$ , the industry code with the maximum employment is chosen.

Both BLS and the Census Bureau have strong incentives and extensive statistical programs to assign detailed and accurate industry codes at the establishment-level. For BLS, the QCEW program yields high quality industry codes from the Annual Refiling Survey as well as BLS business surveys. For the Census Bureau, periodic business surveys and the Economic Census provide rich sources of information on industry. BLS also shares their industry codes with the Census Bureau. The Census Bureau also obtains codes from SSA as part of the first step of identifying new businesses. The SSA industry code uses the information provided in the application for a new EIN (Form SS-4). While SSA industry codes are a useful first step, the Census Bureau has a clear hierarchy for industry codes in their business register and their business statistical programs, with the Economic Census (and related surveys) and BLS codes preferred (see Walker 1997).

In complementary work, Haltiwanger and Spletzer (2020) show that the fraction of the variance of earnings accounted for by industry effects is very similar using either BLS or census codes but is much smaller using the industry codes census obtains from SSA. Moreover, Bloom et al. (2018) indicate that the same SSA

<sup>10</sup>Details of our dataset construction can be found in online Appendix G.

<sup>11</sup>The level of industry aggregation trades off tractability versus comprehensiveness. Note that four-digit NAICS industries aggregate six-digit industries into “NAICS Industry Groups,” which for ease of exposition, we refer to simply as “industries.” Haltiwanger and Spletzer (2022) measure rising inequality at different levels of NAICS aggregation, and demonstrate that the vast majority of rising between-industry inequality occurs at the four-digit NAICS level.

microdata used in Song et al. (2019) has missing industry codes for all new firms post 2002. Table 2 of Bloom et al. (2018) shows that the employment share of EINs with missing industry codes increased from 4 percent in 1980–1986 to 24 percent in 2007–2013 in their microdata. Our inference is that the high-quality industry codes from BLS and the Census Bureau yield a more accurate characterization of the role of industry variation in accounting for earnings dispersion. In the analysis that follows, the LEHD, CPS-LEHD, and OEWS data and analysis are all based on the QCEW sample frame with consistent industry codes. The LBD analysis uses the Census Bureau’s business register frame which differs from the QCEW frame.<sup>12</sup>

## II. Earnings Inequality within and between Firms and Industries

Unless otherwise stated, the empirical analysis uses the LEHD data infrastructure. When we turn to the CPS-LEHD, OEWS, and LBD analysis, we are explicit about the use of those data sources.

Letting  $i$  index the worker,  $j$  the firm,  $k$  the industry, and  $t$  the year, we can write the variance of real annual log earnings  $y$  as

$$\begin{aligned}
 (1) \quad \underbrace{\text{var}[y_t^{i,j,k,p} - \bar{y}^p]}_{\text{total dispersion}} &= \underbrace{\text{var}[y_t^{i,j,k,p} - \bar{y}^{j,k,p}]}_{\text{within-firm}} + \underbrace{\text{var}[\bar{y}^{j,k,p} - \bar{y}^p]}_{\text{between-firm}} \\
 &= \underbrace{\text{var}(y_t^{i,j,k,p} - \bar{y}^{j,k,p})}_{\text{within-firm}} + \underbrace{\text{var}(\bar{y}^{j,k,p} - \bar{y}^{k,p})}_{\substack{\text{between-firm,} \\ \text{within-industry}}} + \underbrace{\text{var}(\bar{y}^{k,p} - \bar{y}^p)}_{\text{between-industry}}
 \end{aligned}$$

We estimate this variance decomposition separately by seven-year intervals denoted by  $p$ . Note that, for a given interval  $p$ , average earnings are represented by  $\bar{y}^p$  for all workers,  $\bar{y}^{j,k,p}$  for firm  $j$ , and  $\bar{y}^{k,p}$  for industry  $k$ . Table 1 shows that for all workers, the variance of earnings increases from 0.794 in the first interval (1996–2002) to 0.915 in the third interval (2012–2018). Of this 0.121 increase, 0.018 (14.9 percent) occurs within firms, 0.028 (23.1 percent) between firms but within industries, and 0.075 (61.9 percent) between industries. These estimates state that between-industry variance growth accounts for 72.8 percent ( $= 0.075 / (0.028 + 0.075)$ ) of the between-firm contribution to increasing inequality.<sup>13</sup>

It is important to distinguish between a cross-sectional variance decomposition versus a growth decomposition. At a given point in time, the majority (58.0 percent to 64.6 percent) of variance is within firms. However, this within-firm person component of earnings variance is becoming less important over time. Growth in the within-industry firm component is positive but much smaller than between-industry

<sup>12</sup>Online Appendix H.2 shows that the patterns of changing industry dispersion largely overlap between the LBD and LEHD.

<sup>13</sup>We use the LBD to consider the restrictions to 18 states as well as the restrictions to firms with more than 20 employees. In addition, we consider the sensitivity to using the EIN-based firm as the definition of the business compared to using an establishment. Results are robust to these variations, see online Appendix Table H1.

TABLE 1—VARIANCE DECOMPOSITION, BY SEVEN-YEAR INTERVAL

	Interval 1: 1996–2002 (1)	Interval 2: 2004–2010 (2)	Interval 3: 2012–2018 (3)	Growth: 1 to 3 (4)
<i>Panel A. Variance, in levels</i>				
Total variance	0.794	0.862	0.915	0.121
Within-firm	0.512	0.532	0.531	0.018
Between-firm, within-industry	0.112	0.127	0.140	0.028
Between-industry	0.170	0.203	0.245	0.075
<i>Panel B. Variance, as percent of total</i>				
Within-firm	64.6	61.7	58.0	14.9
Between-firm, within-industry	14.0	14.7	15.3	23.1
Between-industry	21.4	23.6	26.8	61.9
<i>Panel C. Other measures</i>				
Sample size (millions)	239.4	249.2	269.7	
Number of firms (thousands)	470	460	466	
Number of NAICS industries	301	301	301	

Notes: Authors’ tabulations of LEHD microdata. Tabulations include workers with annual real earnings > \$3,770 in EINs with 20 or more employees. See equation (1) for definitions.

growth. The between-industry component grows substantially over time, from 21.4 percent in the first interval to 26.8 percent in the third interval.<sup>14</sup>

### III. The Industries That Drive Increasing Inequality

We have demonstrated that almost two-thirds of the growth in inequality occurs between rather than within industries. We now propose a measure of a particular industry’s contribution to inequality and assess how this varies across industries. Between-industry variance growth can be expressed as

$$(2) \quad \underbrace{\Delta \text{var} [\bar{y}^{k,p} - \bar{y}^p]}_{\text{between-industry variance growth}} = \sum_{k=1}^{301} \underbrace{\Delta \left( \frac{N^{k,p}}{N^p} \right)}_{\text{employment share}} \underbrace{(\bar{y}^{k,p} - \bar{y}^p)^2}_{\text{relative earnings}} ,$$

industry *k*’s contribution to between-industry variance growth

where *N* counts worker-employer-year combinations (i.e., employment), *N<sup>k,p</sup>* is total employment in industry *k* in interval *p*, and *N<sup>p</sup>* is total employment in interval *p*. We define industry *k*’s contribution to between-industry variance growth as  $\Delta \left( \frac{N^{k,p}}{N^p} \right) (\bar{y}^{k,p} - \bar{y}^p)^2$ .

There are a total of 301 four-digit NAICS industries in our LEHD data. The distribution of contributions and employment shares by industry are depicted in Figure 1. Industries are rank ordered in the contribution to the change in between-industry earnings dispersion from the least to greatest. Strikingly, this distribution is highly skewed

<sup>14</sup>Online Appendix B shows that the rising between-industry contribution occurs throughout the distribution of earnings.



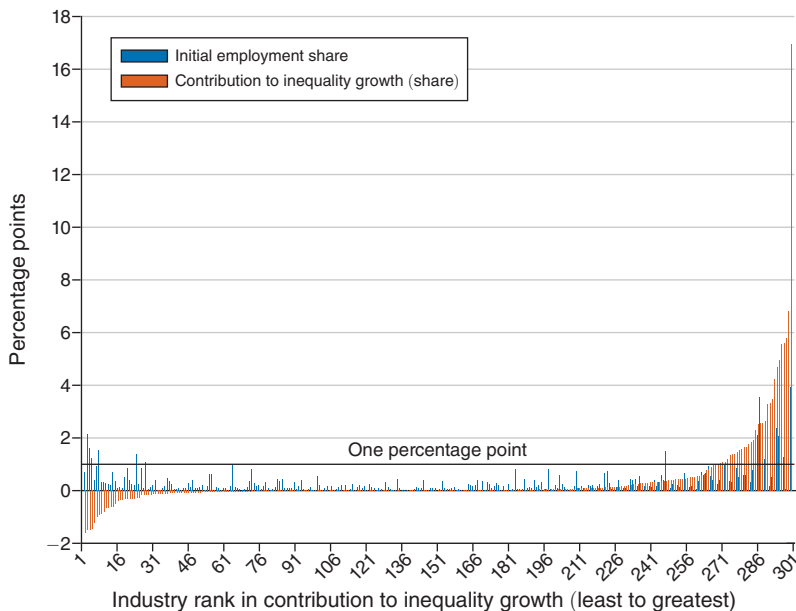


FIGURE 1. INDUSTRY CONTRIBUTION TO BETWEEN-INDUSTRY INEQUALITY AND INITIAL EMPLOYMENT SHARE

Notes: Authors' tabulations of LEHD microdata. Tabulations include workers with annual real earnings  $> \$3,770$  in EINs with 20 or more employees. See equation (2) for definitions.

to the right.<sup>15</sup> There are thirty industries in the right tail of Figure 1 that account for more than 1 percent of inequality growth and five industries in the left tail that account for less than  $-1$  percent. These industries account for 98.1 percent and  $-7.2$  percent of the overall contribution, respectively. The top 30 industries account for almost 40 percent of employment while the bottom 5 industries account for about 5 percent of employment.<sup>16</sup>

To summarize these patterns, Table 2 groups industries by their contributions to increasing inequality. There are five industries that each contribute more than 5 percent of between-industry variance growth, accounting for 40.7 percent of between-industry variance growth. These five industries have 8.8 percent of total employment. An additional 25 industries each contribute between 1 percent and 5 percent of between-industry variance growth, accounting for 57.4 percent of

<sup>15</sup> See online Appendix Table A7 for summary statistics of this distribution.

<sup>16</sup> We also present evidence (see online Appendix H.2) that rising between-industry dispersion played an important role in earlier decades (back to 1980). There is considerable overlap but also distinct differences in the nature of the between-industry contribution in earlier decades. Similar to our main sample period, the contribution is right skewed with industries such as Restaurants and Other Eating Places (7225) and Software Publishers (5112) being important contributors even in this early period. However, some industries that were almost nonexistent (e.g., industries where internet portals and search engines are located) not surprisingly played little role in the earlier period. Moreover, manufacturing played more of a drag on rising inequality in the earlier period with key high-paying manufacturing industries such as Motor Vehicle Parts Manufacturing (3363) and Iron and Steel Mills and Ferroalloy Manufacturing (3311) exhibiting substantial declines in employment in the earlier period.

TABLE 2—INDUSTRY CONTRIBUTIONS TO BETWEEN-INDUSTRY VARIANCE GROWTH, BY VARIANCE CONTRIBUTION

Industry share of between-industry variance growth	Number of industries	Total employment share (%) (1)	Total contribution to between-industry variance growth (2)	Total share of between-industry variance growth (%) (3)
> 5%	5 industries	8.8	0.031	40.7
1% to 5%	25 industries	30.5	0.043	57.4
0.05% to 1%	71 industries	21.8	0.017	22.3
−0.05% to 0.05%	145 industries	19.3	−0.000	−0.1
< −0.05%	55 industries	19.7	−0.015	−20.3
Overall	301 industries	100.0	0.075	100.0

Notes: Authors' tabulations of LEHD microdata. Tabulations include workers with annual real earnings > \$3,770 in EINs with 20 or more employees. Employment shares are calculated as the average of 1996–2002 and 2012–2018 employment shares. See equation (2) for definitions.

between-industry variance growth. In total, the top 30 industries—about 10 percent of all four-digit NAICS industries—account for 98.1 percent of between-industry variance growth and 39.2 percent of employment.

As nearly two-thirds of the growth in US earnings dispersion has occurred between rather than within industries, these thirty industries account for most of increasing inequality. We provide detail about these thirty industries in Table 3 (sorted by NAICS). The largest contribution is from Restaurants and Other Eating Places (7225), which alone accounts for 16.9 percent of between-industry variance growth. The second-largest contribution occurs among Other General Merchandise Stores (4529), which accounts for 6.8 percent. While the most important two industries to increasing inequality tend to offer low-paying jobs, the other three industries that account for more than 5 percent of between-industry variance growth are high-paying: Software Publishers (5112), Computer Systems Design and Related Services (5415), and Management of Companies (5511).

The top 30 industries reflect a small number of industry clusters that are notable for undergoing structural transformations that have been the subject of independent analysis. Eleven of the 19 high-paying industries have been defined as high-tech industries in terms of STEM intensity by Hecker (2005). These innovative industries in combination account for about one-third of the between-industry increase in earnings dispersion. As discussed in Oliner, Sichel, and Stiroh (2007) and Fernald (2014), these industries are characterized as the source of rapid technological advances in information and communications technologies. The transformation of the retail sector accounts for another one-third of the increase. These include the large industries in retail that have undergone the most significant business model transformations towards large national chains (see Foster, Haltiwanger, and Krizan 2006). High-paying Financial Services and Health Care Sectors also play an important role. For the latter, there are also low-paying health care sectors (e.g., retirement centers) making a substantial contribution. Outsourcing is connected to the top 30 industries. As documented by Dey, Houseman, and Polivka (2010) and Dorn et al. (2023), occupations such as protective service,

TABLE 3—INDUSTRY CONTRIBUTIONS TO BETWEEN-INDUSTRY VARIANCE GROWTH, TOP 30 INDUSTRIES

NAICS	Industry title	Employment share percent		Relative earnings		Share of between industry variance growth (%)
		Average	Change	Average	Change	
		(1)	(2)	(3)	(4)	(5)
2111	Oil and Gas Extraction	0.3	-0.0	1.012	0.247	1.8
2131	Support Activities for Mining	0.5	0.3	0.374	0.191	1.4
3254	Pharmaceutical Manufacturing	0.5	-0.1	0.799	0.203	1.6
3344	Semiconductor Manufacturing	0.8	-0.5	0.556	0.299	1.4
4234	Professional Equipment Wholesaler	0.7	-0.0	0.557	0.190	1.9
<b>4441</b>	<b>Building Material and Supplies</b>	0.9	0.1	-0.293	-0.180	1.5
<b>4451</b>	<b>Grocery Stores</b>	2.4	0.0	-0.378	-0.194	4.7
<b>4481</b>	<b>Clothing Stores</b>	0.7	-0.0	-0.607	-0.244	2.6
<b>4529</b>	<b>Other General Merchandise Stores</b>	1.4	1.5	-0.539	-0.051	6.8
5112	Software Publishers	0.5	0.2	1.009	0.186	5.6
5182	Data Processing Services	0.3	-0.0	0.545	0.301	1.3
5191	Other Information Services	0.2	0.3	0.798	0.699	5.8
5221	Depository Credit Intermediate.	2.1	0.0	0.189	0.234	2.5
5231	Securities Brokerage	0.5	-0.1	0.866	0.204	1.1
5239	Other Financial Investment Activity	0.3	0.1	0.834	0.388	3.3
5241	Insurance Carriers	1.6	-0.4	0.488	0.167	2.3
5413	Archit. and Engineering Services	1.2	0.1	0.469	0.161	2.6
5415	Computer Systems Design	1.7	0.9	0.663	0.012	5.6
5416	Mgmt. and Scientific Services	0.9	0.6	0.381	0.069	1.8
5417	Scientific Research Services	0.8	-0.1	0.741	0.244	3.3
5511	Management of Companies	2.0	-0.1	0.471	0.201	5.0
<b>5613</b>	<b>Employment Services</b>	3.9	0.6	-0.685	0.017	2.5
<b>5617</b>	<b>Services to Buildings and Dwell</b>	1.1	0.3	-0.493	-0.002	1.1
6211	Offices of Physicians	1.7	0.5	0.254	0.099	1.6
<b>6216</b>	<b>Home Health Care Services</b>	0.8	0.4	-0.525	-0.016	1.7
6221	General Medical and Hospitals	4.5	0.5	0.205	0.162	4.2
6233	Continuing Care Retirement	0.6	0.4	-0.493	-0.001	1.2
<b>6241</b>	<b>Individual and Family Services</b>	0.8	0.6	-0.490	-0.155	3.5
<b>7139</b>	<b>Othr. Amusement and Recreation</b>	0.6	0.1	-0.594	-0.106	1.7
<b>7225</b>	<b>Restaurants and Othr. Eat Places</b>	4.9	2.0	-0.739	-0.027	16.9

Notes: Authors' tabulations of LEHD microdata. Tabulations include workers with annual real earnings >\$3,770 in EINs with 20 or more employees. Average log earnings for industry  $k$  are relative to the economy average. The 1996–2002 and 2012–2018 intervals are averaged. Changes are the growth (or decline) from 1996–2002 to 2012–2018. See equation (2) for definitions. Top 11 low-paying industries in **bold**.

transportation and material moving, building and grounds cleaning and maintenance, and food preparation and serving are increasingly employed in industries that provide services to other firms. Important industries in the top 11 low-paying industries that fit this description include Employment Services (5613) and Services to Buildings and Dwellings (5617). As we will see below using the OEWS, we find patterns consistent with the shift of such occupations away from the top 19 high-paying industries.

What about the other 271 four-digit NAICS industries? Figure 1 highlights they make relatively little contribution. Using Table 2, there are 145 industries that each

TABLE 4—INDUSTRY CONTRIBUTIONS TO BETWEEN-INDUSTRY VARIANCE GROWTH, BY AVERAGE EARNINGS

Industry relative earnings	Number of Industries	Total employment share (1)	Total contribution to between-industry growth (2)	Total share of between-industry growth (3)	Shift-share	
					Employment (4)	Earnings (5)
Overall	301	100.0%	0.075	100.0%	14.0%	86.0%
<i>Panel A. 30 industries with variance contribution &gt; 1%</i>						
High-paying	19	21.1%	0.041	54.1%	16.1%	83.9%
Low-paying	11	18.1%	0.033	44.1%	68.3%	31.7%
<i>Panel B. 271 industries with variance contribution ≤ 1%</i>						
High-paying	146	34.9%	0.001	1.3%		
Low-paying	125	25.9%	0.000	0.6%		

Notes: Authors' tabulations of LEHD microdata. Tabulations include workers with annual real earnings > \$3,770 in EINs with 20 or more employees. Employment shares are calculated as the average of 1996–2002 and 2012–2018 employment shares. Industry  $k$ 's contribution to between-industry variance growth is specified in equation (2). The shift-share calculations for changing employment and earnings follow equation (3). Shift-share results are summed across industries and normalized by the total contribution so that the two components sum to 100 percent. The two rows for the 271 industries with variance contribution ≤ 1 percent have missing cells because the denominator for the shift-share decomposition is close to zero.

contribute approximately 0.0 percent (to be precise, greater than  $-0.05$  percent and less than 0.05 percent) to between-industry variance growth. This says that almost one-half of all four-digit NAICS industries contribute essentially nothing to inequality growth. There are 71 industries that contribute between 0.05 percent and 1.0 percent, accounting for 22.3 percent of between-industry variance growth. These industries are basically offset by another 55 industries that have a negative contribution ( $< -0.05$  percent), accounting for  $-20.3$  percent of between-industry variance growth.

As seen in Table 4, the top 30 industries include 19 high-paying industries that account for 54.1 percent of between-industry variance growth, and 11 low-paying industries that account for 44.1 percent of between-industry variance growth. The other 271 industries that have small contributing and offsetting contributions to increasing inequality do not occur systematically among high-paying versus low-paying industries. 146 high-paying industries account for 1.3 percent of between-industry variance growth, and 125 low-paying industries account for only 0.6 percent of between-industry variance growth.

Changes in earnings and employment share determine an industry's contribution to growth in inequality. This is seen in the expression defining industry  $k$ 's contribution to between-industry variance growth:  $\Delta\left(\frac{N^{k,p}}{N^p}\right)(\bar{y}^{k,p} - \bar{y}^p)^2$ . If an industry with relatively high earnings exhibits an earnings increase, then, *ceteris paribus*, inequality will increase. Analogously, inequality will increase if an industry with relatively low earnings exhibits an earnings decrease. In contrast, when average earnings in an industry moves closer to the overall average, inequality decreases.

Employment shares also determine industry-level contributions to inequality. An industry's earnings changes will have larger effects on inequality when its employment share is larger. Changes in an industry's employment share will have smaller effects on inequality when that industry's pay is more similar to the overall average.

Employment gains among very high- and very low-paying industries tend to increase inequality.

In Table 4, we report the relative importance of earnings changes versus employment changes using a shift share decomposition. Industry  $k$ 's contribution to between-industry variance growth is  $\Delta\left(\frac{N^{k,p}}{N^p}\right)(\bar{y}^{k,p} - \bar{y}^p)^2$ . We can use a standard shift-share decomposition to express this change in terms of the components attributable to changes in employment versus earnings:

$$(3) \quad \underbrace{\Delta\left(\frac{N^{k,p}}{N^p}\right)(\bar{y}^{k,p} - \bar{y}^p)^2}_{\text{industry } k\text{'s contribution to between-industry variance growth}} = \underbrace{\overline{(\bar{y}^{k,p} - \bar{y}^p)^2} \Delta\left(\frac{N^{k,p}}{N^p}\right)}_{\text{shift-share: employment}} + \underbrace{\frac{\overline{N^{k,p}}}{N^p} \Delta(\bar{y}^{k,p} - \bar{y}^p)^2}_{\text{shift-share: earnings}},$$

where  $\overline{(\bar{y}^{k,p} - \bar{y}^p)^2}$  and  $\frac{\overline{N^{k,p}}}{N^p}$  are averages of intervals 1 and 3. We do this for our top 30 industries, distinguished by high-paying and low-paying industries (we do not present the shift share estimates for the other 271 industries since the denominator of the shift share is very close to zero). Among the 19 high-paying industries, 83.9 percent of between-industry variance growth is accounted for by changing relative earnings, and the remaining 16.1 percent is accounted for by changing employment shares. Among the 11 low-paying industries, the relative importance of earnings versus employment is reversed: 68.3 percent of between-industry variance growth is accounted for by changing employment shares, and the remaining 31.7 percent is accounted for by changing relative earnings. These results highlight different explanations for why between-industry variance growth is increasing at the opposite tails of the earnings distribution. Inequality growth at the top of the earnings distribution is a story of increasing earnings, whereas inequality growth at the bottom of the earnings distribution is a story of increasing employment.

These two different explanations for increasing inequality among low- versus high-paying industries is evident in the earnings and employment changes of the thirty industries listed in Table 3. All of the 19 high-paying industries exhibit earnings increases during our time period. The most rapid growth is found in Other Information Services (5191), which had a 69.9 log point (101.2 percent) increase in relative earnings.<sup>17</sup> Of the remaining high-paying industries, nine had earnings increases in excess of 20 log points (22.1 percent), six had increases between 10 (10.5 percent) and 20 log points, and three had increases less than 10 log points.

Most of the 11 low-paying industries exhibit earnings decreases, yet they are smaller in absolute value than the earnings increases among the high-paying industries. The only low-paying industry with a decline greater in magnitude than 20 log points (22.1 percent) is Clothing Stores (4481), which had a 24.4 log point (27.6 percent) decrease in relative earnings. Of the remaining low-paying industries, four had earnings declines between 10 (10.5 percent) and 20 log points, and five had

<sup>17</sup> We convert log differentials to proportionate changes using the expression  $e^x - 1$ . For small differences, log points are approximately equal to the percentage change.

earnings declines between 0 and 10 log points. One industry, Employment Services (5613), exhibited a relatively small increase in earnings.

On the other hand, changes in employment are more important for the 11 low-paying industries than for the 19 high-paying industries. Two low-paying industries in Table 3 stand out: Restaurants and Other Eating Places (7225) had a 2.0 percentage point increase in employment share, and Other General Merchandise Stores (4529) had a 1.5 percentage point increase in employment share. Eight of the other low-paying industries have smaller employment share increases (less than one percentage point), and one industry (Clothing Stores, 4481) had a declining employment share. Among the 19 high-paying industries, none had employment share increases exceeding 1 percentage point, ten had small employment share increases (less than 1 percentage point), and about one-half (9) of the high-paying industries had declining employment shares.

In the analysis that follows, we also use the CPS-LEHD integrated data, the OEWS and LBD data to provide further insights into the role of rising between-industry dispersion. It is worth highlighting that all of these alternative sources provide a similar quantification of the contribution of the top 30 industries listed above to rising between-industry dispersion. While more detail is provided below, the share of the between-industry increase in dispersion from 1996–2002 to 2012–2018 accounted for by the top 30 industries is 98.1 percent in the LEHD data, 105.5 percent in the CPS-LEHD data, 96.2 percent in the OEWS data, and 94 percent in the LBD data.

#### IV. Firm and Worker Composition in the Top 30 Industries

##### A. *Mega Firms*

Changes in the employment shares and size-earnings premia for mega (10,000+) firms play a critical role in accounting for rising between-industry earnings inequality. Table 5 shows descriptive statistics of employment and earnings in mega firms and non-mega firms in our four industry groups. One immediate result in Table 5 is that employment has shifted over time to the top 30 industries. The employment share of the top 30 industries increased by 8.2 percentage points, with most of this increase (6.0 percentage points) among the 11 low-paying industries. The employment share of the other 271 industries analogously declined by 8.2 percentage points, with most of this decline (6.8 percentage points) among the 146 high-paying industries.

The substantial increase in the employment share of the top 30 industries is driven by mega firms. This is evident in both Table 5 and Figure 2. Figure 2 shows the change in employment share by detailed size class for each of our four industry groups.<sup>18</sup> The employment share of the 11 low-paying industries increased in every size class, with mega firms exhibiting the largest increase (2.5 percentage points). The 19 high-paying industries had a smaller increase in employment, but most of this increase (1.4 of a total of 2.2 percentage points) is accounted for by mega firms. Given the high average relative pay of mega firms in the high-paying

<sup>18</sup>The corresponding employment share levels in the first interval (1996–2002) and in the third interval (2012–2018) are given in online Appendix Figure F1.

TABLE 5—CHANGES IN EMPLOYMENT AND EARNINGS, BY INDUSTRY EARNINGS, MEGA FIRMS VERSUS OTHERS

Industry relative earnings	Number of industries	Firm employment	Employment share		Relative earnings	
			Average (1)	Change (2)	Average (3)	Change (4)
<i>Panel A. 30 industries with variance contribution &gt; 1%</i>						
High-paying	19 industries	Any	21.1%	2.2%	0.440	0.177
		10,000+	3.8%	1.4%	0.579	0.145
		< 10,000	17.3%	0.8%	0.410	0.174
Low-paying	11 industries	Any	18.1%	6.0%	-0.586	-0.069
		10,000+	4.3%	2.5%	-0.492	-0.125
		< 10,000	13.8%	3.5%	-0.613	-0.061
<i>Panel B. 271 industries with variance contribution ≤ 1%</i>						
High-paying	146 industries	Any	34.9%	-6.8%	0.281	0.046
		10,000+	3.9%	-1.2%	0.646	0.042
		< 10,000	31.0%	-5.7%	0.236	0.052
Low-paying	125 industries	Any	25.9%	-1.3%	-0.325	-0.002
		10,000+	3.3%	-0.5%	-0.404	-0.061
		< 10,000	22.6%	-0.9%	-0.314	0.006

Notes: Authors' tabulations of LEHD microdata. Tabulations include workers with annual real earnings > \$3,770 in EINs with 20 or more employees. Averages and changes use the employment shares and earnings from the 1996–2002 and 2012–2018 intervals. Average log earnings are relative to the economy average.

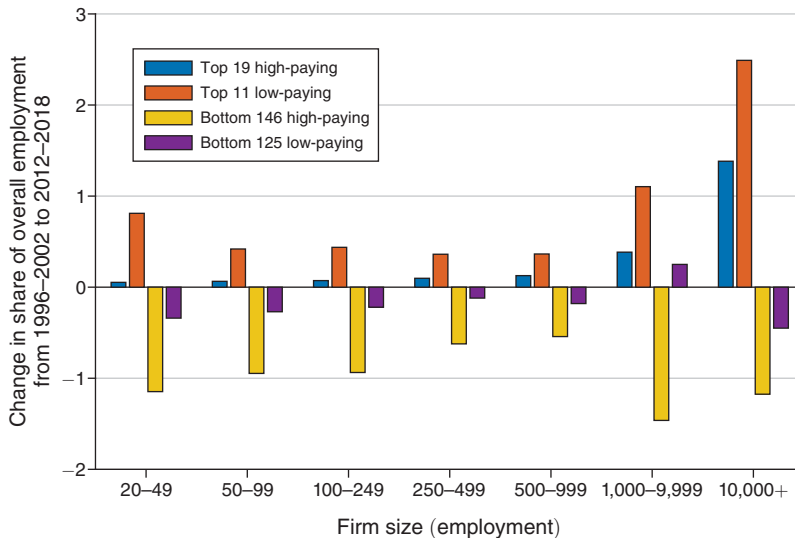


FIGURE 2. CHANGE IN EMPLOYMENT SHARE BY SIZE CLASS, BY INDUSTRY GROUP

Notes: Authors' tabulations of linked LEHD microdata. Tabulations include workers with annual real earnings > \$3,770 in EINs with 20 or more employees. Changes in the employment shares are expressed in terms of percentage points. The denominator is total employment across all size classes and industry groups.

industries (57.9 log points, or 77.9 percent) and the low pay of mega firms (-49.2 log points, or -63.6 percent) in the low-paying industries, these shifts

in employment to mega firms contributed to rising between-industry earnings inequality.<sup>19</sup>

Mega firms also play a key role in the changing earnings of the top 30 industries. For the 11 low-paying industries, the relative pay of mega firms decreased by 12.5 log points (13.3 percent) compared to a decline of 6.1 log points (6.3 percent) for the non-mega firms. Both mega firms and non-mega firms in the 19 high-paying industries exhibit large earnings increases: 14.5 log points (15.6 percent) for mega firms and 17.4 log points (19.0 percent) for non-mega firms. Earnings at mega firms increased relative to the smallest firms in the top-paying industries but not by as much as the increase in relative earnings at large but not mega firms.<sup>20</sup> In contrast, relative earnings increases at mega firms in the 146 remaining high-paying industries are modest (4.2 log points, or 4.3 percent) compared to 14.5 log points in the top 19 high-paying industries. Similarly, relative earnings declines at the mega firms in the remaining 125 low-paying industries are modest (−6.1 log points, or −6.3 percent) compared to the −12.5 log points in the top 11 low-paying industries.

### B. Education and Occupation

To shed light on the changing education composition of the top and bottom industries, we turn to the CPS-LEHD integrated data. Figure 3 shows the change in employment in the top 30 industries from 1996–2002 to 2012–2018. Both low- and high-paying industries had increases in the educational attainment of the workers that they employ. These changes were much more dramatic in the top 19 high-paying industries. The share of workers with bachelor's degrees at these high-paying industries increased by 7.0 percent, workers with advanced degree increased by 8.4 percent, and workers with a high school diploma declined by 8.7 percent.<sup>21</sup>

For occupation, we turn to the OEWS published data.<sup>22</sup> In Figure 4, we consider employment changes across all 22 occupation groups in the top 30 industries. The occupation groups are ranked from left to right by the changes in employment shares in the top 30 industries. There are substantial differences in the changing mix of occupations across the top 19 and bottom 11. The top 19 industries have large increases in Business and Financial Operations (13) and Computer and Mathematical Science (15) with accompanying large declines in Office and

<sup>19</sup>Online Appendix Table H2 shows that the patterns of changes in employment shares by mega firms is robust to using the national and enterprise concepts available in the Business Dynamic Statistics (BDS). The BDS is derived from the LBD.

<sup>20</sup>Online Appendix Figure F3 shows the cross-sectional size-earnings premia for the 1996–2002 and the 2012–2018 intervals. Among the top 19 high-paying industries, the size-earnings profile shifts upward, with increases in all size classes. Online Appendix F also provides more details of the changing patterns by firm size.

<sup>21</sup>Online Appendix Figure D1 shows the distribution of employment by education group in the top 19 high-paying and 11 low-paying industries for the 1996–2002 and 2012–2018 periods. In the initial period, the top 19 high-paying industries had significantly higher shares of workers with a bachelor's or advanced degree while the top 11 low-paying industries had higher shares of workers with only a high school diploma or less. These differences became much more pronounced by 2012–2018 with sharp increases in the share of workers with bachelor's and advanced degrees in the top 19 high-paying industries and accompanying declines in the share of workers with a high school diploma or less.

<sup>22</sup>Online Appendix Figure G1 shows that the CPS-LEHD yields similar patterns on changes in occupations by industry.



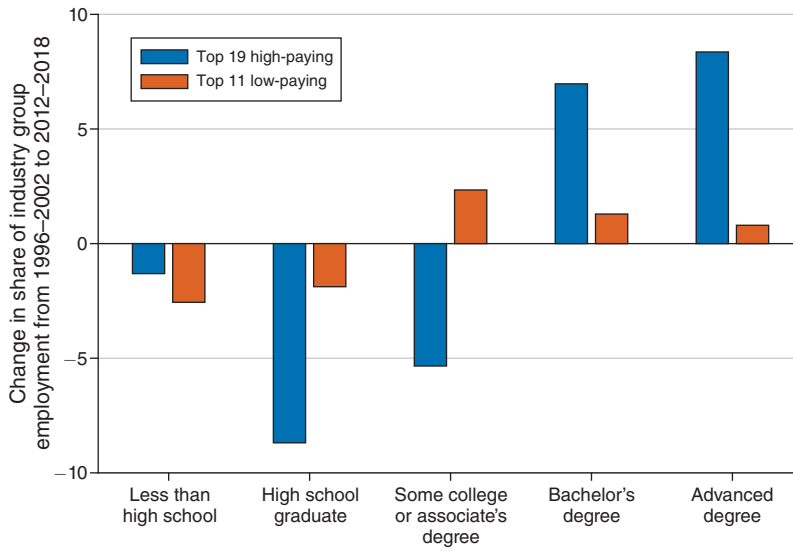


FIGURE 3. CHANGE IN EMPLOYMENT SHARE BY EDUCATIONAL ATTAINMENT, BY INDUSTRY GROUP

Notes: Authors' tabulations of linked CPS-LEHD microdata. Tabulations include workers with annual real earnings > \$3,770. Changes in the employment shares are expressed in terms of percentage points. The denominator is total employment in the respective industry group.

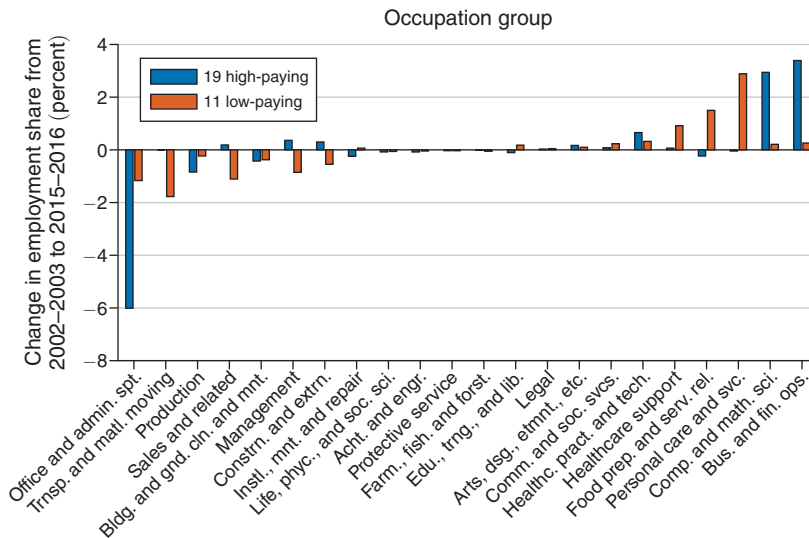


FIGURE 4. CHANGE IN EMPLOYMENT BY OCCUPATION AND INDUSTRY GROUP

Note: Authors' calculations of published OEWS aggregates.

Administrative Support (43), a low wage occupation group. These patterns are consistent with the dramatic innovations in information and communication technology largely developed by top-paying industries. Related, much of the role of Office and Administrative Support (43) tasks in these high-paying industries are

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/518071142017007005>