Industries, Mega Firms, and Increasing Inequality

John Haltiwanger, University of Maryland

Henry Hyatt, U.S. Census Bureau

James R. Spletzer, U.S. Census Bureau

May 7, 2021

PRELIMINARY

Abstract

Most of the rise in overall earnings inequality is accounted for by rising between industry inequality from about ten percent of 4-digit NAICS industries. These industries are in the tails of the earnings distribution. For example, high-paying industries in the top ten percent such as Software Publishers have exhibited an increasing size premium along with a rising share of employment at especially the largest (mega) firms. The rising size-earnings premium in these industries is accounted for by both an increase in the average AKM firm premium and the average AKM person effect. Low-paying industries in the top ten percent such as Restaurants and Other Eating Places have exhibited a decline in the size-earnings premium along with a rising share of employment at the mega firms. The declining size premium in these industries is accounted for by both a decrease in the average AKM firm premium and the average AKM person effect. Strikingly, the remaining ninety percent of industries contribute little to between industry earnings inequality and exhibit little change in the employment share at mega firms. We thus find that the rise of mega firms in a relatively small number of industries plays a critical role in rising earnings dispersion across firms and industries. We also find that increased sorting and segregation of workers as well as rising dispersion in average firm premium are important for the rising between industry dispersion in the dominant ten percent of industries. Importantly, it is increased sorting and segregation between industries rather than within industries that matters.

Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed (CBDRB-FY21-046, CBDRB-FY21-143, AND CBDRB-FY21-155). We thank Keith Bailey and Andrew Foote for helpful comments, along with a special thank you to Christina Patterson whose discussant comments led to a substantial improvement of this paper. John Haltiwanger was also a Schedule A employee of the U.S. Census Bureau at the time of the writing of this paper.

I. Introduction

A growing number of studies attribute increases in earnings inequality to rising between-firm dispersion.[1] We confirm this pattern with comprehensive matched employer-employee data from 1996 to 2018. Our contribution is to explore and emphasize the dominant role of industry effects in accounting for rising between-firm dispersion.[2] Rising between-industry dispersion accounts for most of the overall increase in earnings inequality. Rising between-industry dispersion is driven by a relatively small number of industries. About ten percent of 4-digit NAICS industries account for virtually all of the increase in between-industry dispersion, while accounting for less than 40% of employment.[3] These industries are in the tails of the earnings distribution including high-paying industries such as software publishing and low-paying industries such as restaurants and other eating places. Remarkably, the remaining 90 percent of 4-digit industries individually contribute little to rising between industry earnings inequality.

We provide further insights about rising between-industry inequality using an Abowd, Kramarz, and Margolis (1999) (AKM) decomposition of earnings. Changing composition of workers across industries through sorting (high wage workers are more likely to work in

---

[1] Barth et al. (2016) and Song et al. (2019) provide evidence for the U.S. 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, Foster, and Haltiwanger (2004). Card, Heining, and Kline (2013) consider the role of firms in rising inequality in Germany, and Card, Cardoso, and Kline (2016) consider evidence from Portugal.

[2] Haltiwanger and Spletzer (2020a, 2020b) use a closely related data infrastructure and also emphasize the dominant contribution of rising between-industry dispersion. However, each of these papers proceeds in quite distinct directions from this common starting point. The first paper documents that the rising between-industry dispersion is closely linked to changing occupational differentials and occupational mix across industries. The second paper documents that the rising dispersion across industries accompanied by declining labor market fluidity implies that the rungs of the job ladder have become further apart, and it is more difficult for a worker to get on and climb the job ladder. These earlier papers do not use the AKM decomposition to shed light on the nature of the rising between-industry dispersion. Moreover, the current paper is distinct in documenting and analyzing that a small fraction of industries account for virtually all the rising between-industry dispersion.

[3] For males, 28 industries individually account for 1% or more of rising between industry dispersion and cumulatively account for 99% of the increase in between industry dispersion. These industries account for 32% of male employment. For females, 28 industries (23 overlap with males) individually account for 1% or more of rising between industry dispersion and cumulatively account for 97% of the increase in between industry dispersion. These industries account for 43% of female employment. Restaurants and other eating places is the top contributing industry for both males and females (18.8% for males and 14.2% for females).

industries with high average firm effects) and segregation (high wage workers are more likely to work together in the same industry) account for most but not all of the industry effects. Importantly, it is increased sorting and segregation between industries, rather than between firms within industries, that primarily matters for rising individual earnings dispersion.

We find some differences in the role of sorting vs. segregation based on whether the dominant ten percent of industries tend to be low-paying vs. high-paying. For low-paying dominant industries, sorting plays an especially important role, as the lowest-paying industries are especially likely to employ low person effect workers with very low firm effects. Rising dispersion through both increased sorting and segregation are about equally important among the high-paying dominant industries. Rising dispersion in between-industry average firm premia also play an important role.[4]

There are also distinct differences across the high and low-paying dominant industries in the changes in the size-earnings premium. In the dominant high-paying industries, we find a rising size-earnings premium. This rising earnings premium is accompanied by a sharp increase in employment share in these dominant high-paying industries, particularly at the largest (mega) firms. In the dominant low-paying industries, we find a decline in the size-earnings premium, accompanied by a sharp increase in employment at the largest (mega) firms.[5] Strikingly, the remaining ninety percent of industries that contribute little to rising between industry inequality exhibit little change in the share of employment at mega firms. We thus find that the rise of mega firms in a relatively small number of industries plays a critical role in rising earnings dispersion across firms and industries.

[4] Our finding of an important role for rising firm premia contrasts with the findings of Song et al. (2019). As we discuss below, this is mostly due to differences in sample period.
[5] The decline yields a flattening of the relationship between earnings and size. Large firms still pay more relative to small firms but the gap declines.

Our findings build on the recent literature that highlights the dominant role of rising between-firm inequality. Our results are closest to those in the recent pathbreaking work of Song et. al. (2019). Using Social Security Administration (SSA) administrative data linking employers and employees, they find a dominant role for rising between-firm earnings inequality. Moreover, using an AKM decomposition, they attribute most of this to changing composition from increasing sorting and segregation of workers across firms. While our results are consistent with these findings, our results depart significantly from theirs in our finding of the dominant role of industries in accounting for rising between-firm inequality. In turn, we find that it is a relatively small share of industries that account for this dominant role of industry effects. In contrast, Song et. al. (2019) present findings that industry effects are unimportant.[6] Our results highlight that the increased sorting and segregation of workers across firms is across a relatively narrow set of industries.

Our findings also add perspective to those in Bloom et. al. (2018) using the same SSA data infrastructure showing a flattening size-earnings premium. We also find a flattening size-earnings premium overall but this masks substantial differences in the changing size-earnings premium across the high-paying and low-paying industries that account for virtually all the increase in earnings inequality. As noted, we find that there is a flattening size-earnings premium in the dominant low-paying industries that account for about half of rising between-industry dispersion in earnings. However, we find a rising size-earnings premium in the dominant high-paying industries that account for the other half of rising between-industry dispersion in earnings. These differences are critical for understanding the contribution of such industries to rising earnings inequality.

[6] See page 22 and Table II of Song et al. (2019).

Song et. al. (2019) and Bloom et. al. (2018) use the Social Security Administration (SSA) data which is comprehensive matched employer-employee data in terms of tracking workers and firms, but, as we discuss below, industry codes in the SSA data are known to be of inferior quality. Our analysis is based on using the comprehensive matched employer-employee data from the Longitudinal Employer-Household Dynamics (LEHD) data infrastructure at Census. While both the SSA and LEHD data are based on high quality (but distinct) administrative matched employer-employee data, the LEHD data infrastructure has high quality detailed industry codes from Bureau of Labor Statistics (BLS) and Census processing. Both statistical agencies have strong incentives to have high-quality, detailed industry codes since the agencies produce key national indicators by detailed industry from their employer data.

The finding that most of the rising earnings inequality is accounted for by a relatively small share of industries changes the narrative and the potential mechanisms driving rising earnings inequality. While we confirm the important role of the changing composition across firms via sorting and segregation, we show that this is driven mostly by changing composition across a relatively small share of industries. This implies that the mechanisms underlying this structural change apply to only a relatively small fraction of industries. Moreover, the small number of dominant industries are characterized by substantial employment increases in mega firms with accompanying changes in the size-earnings premiums.

The paper proceeds as follows. Section II describes the data infrastructure and provides descriptive statistics about the changing distribution of earnings over our sample period from 1996 to 2018. The AKM decomposition methodology is presented in Section III. Section IV presents the results from the AKM decomposition by distinguishing the sorting, segregation, and firm premia contributions. Section V enhances the decomposition to distinguish these firm contributions between and within industries. The dominant contribution of a relatively small share of industries to rising between-industry dispersion is presented in Section VI. The role of the rising importance of mega firms is discussed in Section VII. Concluding remarks are in Section VIII.

II. Data and Descriptive Statistics

II.A The LEHD Data and the Analysis Sample

We use Longitudinal Employer-Household Dynamics (LEHD) linked employer-employee data, which is created by the U.S. Census Bureau as part of the Local Employment Dynamics federal-state partnership. The LEHD data are derived from state-submitted Unemployment Insurance (UI) wage records and Quarterly Census of Employment and Wages (QCEW) data. Every quarter, employers who are subject to state UI laws -- approximately 98% of all private sector employers, plus state and local governments -- are required to submit to the states information on their workers (the wage records, which lists the quarterly earnings of every individual in the firm) and their workplaces (the QCEW, which provides information on the industry and location of each establishment). The wage records and the QCEW data submitted by the states to the U.S. Census Bureau are enhanced with census and survey microdata in order to incorporate information about worker demographics (age, gender, and education) and the firm (firm age and firm size). Abowd et al. (2009) provide a thorough description of the source data and the methodology underlying the LEHD data. A job in the quarterly LEHD data is defined as the presence of an individual-employer match, and earnings is defined as the amount earned from that job during the quarter.

Because states have joined the LEHD program at different times and have provided various amounts of historical data upon joining the LEHD program, the length of the time series of LEHD data varies by state. We use data from the 18 states that have data available from 1996:Q1 through 2018:Q4, which gives us annual data for 23 years.[7] We restrict the LEHD data to jobs in the private sector.

Following Song et.al (2019), we create annual person-level data from the quarterly jobs-level earnings data. We do this as $Y_{it} = \sum_j \sum_{q \in t} Y_{ijq}$, where Y is earnings, "i" is individual, "j" is firm, "q" is quarter, and "t" is year. We use the Federal Employer Identification Number (EIN) as the firm identifier.[8] We follow Abowd, McKinney, & Zhao (2018) and delete any individual with 12 or more jobs during the year. In the annual person-level data created by summing quarterly earnings across all jobs, the firm is defined as the firm that contributes the person's maximum earnings during the year. The annual data has 1,395,000,000 person-year observations (an average of 60,650,000 persons per year).[9]

We create our analytical dataset following the sample restrictions of Song et al. (2019). We restrict to persons aged 20-60 and only keep person-year observations with annual real earnings > \$3770 (=13 weeks * 40 hours per week * \$7.25 minimum wage), with nominal earnings converted to real terms using the 2013=100 PCE deflator. From this sample of 20-60 year olds with annual earnings greater than \$3770, we topcode annual earnings at the 99.999% value (for anyone with earnings in the top 0.001 percent, we replace their earnings with the mean

---

[7] These 18 states are: CA, CO, CT, HI, ID, IL, KS, LA, MD, MN, MT, NC, NJ, OR, RI, TX, WA, and WI. 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.

[8] Haltiwanger and Spletzer (2020b) 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.

[9] All estimates in this paper were prepared following U.S. Census Bureau requirements for disclosure avoidance review. Among other things, the release of results from confidential microdata usually requires estimates be rounded to four significant digits.

earnings of the top 0.001 percent). This annual individual-level data has 1,048,000,000 person-year observations (an average of 45,570,000 persons per year). All of our analysis uses real annual log earnings (y_{it}).

We then define three 7-year intervals (1996-2002, 2004-2010, 2012-2018), reducing the sample to 959,400,000 person-year observations. We estimate interval specific AKM fixed effect regressions (described in the next section) for the largest connected set of males and for the largest connected set of females. We have AKM fixed effects for 938,900,000 person-year observations (an average of 44,710,000 persons per year). Our estimation of AKM fixed effect regressions separately for males and females follows Card, Cardoso, and Kline (2016).

And finally, again following Song et.al (2019), we restrict to firm-year observations with 20 or more persons in the firm. This reduces our sample to 762,900,000 person-year observations. Due to Census disclosure rules, we further restrict the firms with 20 or more employees in each year to have at least one male and one female; this means that the same firms are used for the male and female variance decompositions. The final LEHD data used to create all our results contains 758,400,000 person-year observations (an average of 36,110,000 persons per year). Our analytical sample has 412,500,000 person-year observations for males and 345,900,000 person-year observations for females.

II.B Industry Code

Industry codes play a fundamental role in our analysis. Our basic results use establishment-level industry codes from the BLS QCEW program, aggregated to the Federal EIN level. Aggregation from establishment level data is done using maximum employment (for example, if an EIN has N≥1 establishments with N≥M≥1 industry codes, the industry code with the maximum employment is chosen for the aggregation).

Both BLS and Census 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 the BLS surveys of businesses. For Census, the periodic surveys and the Economic Censuses of businesses provide rich sources of information on industry codes. BLS also shares their industry codes with Census. Census also obtains codes from SSA as part of the first step of identifying new businesses. The industry code from SSA is based on the information provided in the application for a new EIN (the SS-4 form). While SSA industry codes are a useful first step, Census 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 (2020a) 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) indicates that the same SSA micro data used in Song et. al. (2019) has missing industry codes for all new firms post 2002. Their Table 2 shows that EINs with missing industry codes increased from accounting for only 4% of total employment in the 1980-86 period to 24% in 2007-13 in their micro data. Our inference is that the high-quality industry codes from BLS and Census yield a much more accurate characterization of the role of between-industry variation in accounting for earnings dispersion.

II.C Descriptive Statistics

The black line in Figure 1 shows the 1996-2018 time series of the variance of person-level real annual earnings for the full sample including males and females. We see that the variance of earnings is increasing throughout but especially from 1996 to 2012. Figure 1 also shows a decomposition of variance into firm and industry components. Letting "i' index the individual, "j" the firm, and "k" the industry, we can write the variance of real annual log earnings y in year "t" as:

$$var\left(y_t^{i,j,k}-\bar{y}_t\right)=var\left(y_t^{i,j,k}-\bar{y}_t^{j,k}\right)+var\left(\bar{y}_t^{j,k}-\bar{y}_t\right)$$
$$=var\left(y_t^{i,j,k}-\bar{y}_t^{j,k}\right)+var\left(\bar{y}_t^{j,k}-\bar{y}_t^{k}\right)+var\left(\bar{y}_t^{k}-\bar{y}_t\right).$$

The black line in Figure 1 is the variance of earnings $var\left(y_t^{i,j,k}\right)$, the red line is the within-firm variance $var\left(y_t^{i,j,k}-\bar{y}_t^{j,k}\right)$, and the solid blue line is the between-firm variance $var\left(\bar{y}_t^{j,k}-\bar{y}_t\right)$. The between-firm variance is the sum of the between-firm, within-industry variance $var\left(\bar{y}_t^{j,k}-\bar{y}_t^{k}\right)$, the dotted blue line in Figure 1, and the between-industry variance $var\left(\bar{y}_t^{k}-\bar{y}_t\right)$, the dashed blue line. Throughout this paper, industry refers to 301 4-digit NAICS industries.

In Figure 1, the variance of earnings increases from .7548 in 1996 to .9114 in 2018, an increase of roughly 15 log points. 29.2 percent of this increase is within firms (the red line in Figure 1) and 70.8 percent is between firms (the solid blue line in Figure 1). Looking at industry, 20.6 percent of the total increase in variance is within industries and 50.2 percent is between industries. This finding that the between-industry component accounts for more than half of increasing earnings inequality, and 71 percent of between-firm inequality growth, is consistent with our earlier work in Haltiwanger and Spletzer (2020a, 2020b).

Table 1 presents the same variance decomposition using 7-year intervals rather than single years.[10] Table 1 shows the results for all workers the variance of earnings increases from 0.7938 in the first interval (1996-2002) to 0.9152 in the third interval (2012-2018). This 0.1214 increase is decomposed into 0.0181 within firms (14.9 percent), 0.0281 within industries (23.1 percent), and 0.0752 between industries (61.9 percent).

It is important to distinguish between a cross-sectional variance decomposition versus a growth decomposition. At a given point in time, the majority of variance is within firms: 64.6 percent of variance in the first interval is within firms, 61.7 percent in the second interval, and 58.0 percent in the third interval. This declining relative percentage is indicative that the within-firm person component of earnings variance is becoming less important over time. The within-industry firm component of earnings variance increases slightly over time. It is the between-industry component of variance that is growing substantially over time, from 21.4 percent in the first interval to 26.8 percent in the third interval.

Tables 1A and 1B show results of our variance decomposition for males and females, respectively. While results are broadly similar for these two subpopulations, there are some noteworthy differences. Earnings dispersion grew more among females (0.1392) than males (0.1263). For both males and females, consistent with the findings of Song et al. (2019), the vast majority of this increase was across firms rather than within them. However, within-firm dispersion among women accounted for 28.6% of the increase in the variance of earnings growth, while among men this fraction was only 15.5%.

II.D Earnings Percentiles

[10] We follow Song et al. in using 7-year intervals which facilitates the estimation of the AKM earnings decomposition for different time intervals.

The statistics in Table 1 demonstrate that 4-digit NAICS industry accounts for almost two-thirds of the growth of earnings variance. In this section, we present a descriptive analysis to learn where in the earnings distribution industry seems to be more important. We begin by estimating annual earnings for each of the percentiles 1 to 99 for the first (1996-2002) and the third (2012-2018) 7-year intervals, and then calculating the difference between the first and third intervals for each percentile.[11] In our analytical sample, comparing the first and the third intervals, annual earnings declined by more than five log points for the first 34 percentiles, and declined for the first 61 percentiles. However, earnings at the top increased substantially. Earnings in the top 23 percentiles increased by more than five log points, and earnings in the top 13 percentiles increased by more than 10 log points.

We use a simple decomposition to understand how the person, the firm, and the industry help account for the changing distribution of earnings. We can rewrite earnings $y_t^{i,j,k}$ as

$$\left(y_t^{i,j,k} - \bar{y}_t\right) = \left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right) + \left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right) + \left(\bar{y}_t^{k} - \bar{y}_t\right).$$

We estimate each of the terms on the right-hand side for each percentile of $\left(y_t^{i,j,k} - \bar{y}_t\right)$, noting that firm mean earnings $\bar{y}_t^{j,k}$, industry mean earnings $\bar{y}_t^{k}$, and the grand mean of earnings $\bar{y}_t$ are from the full sample of individuals rather than calculated within each percentile. To interpret this exercise, think of individuals in the first percentile, who have earnings between the ½th and 1½th percentiles. We estimate how the earnings of these individuals differ from the earnings of their firm $\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right)$, how the earnings of their firm differ from the earnings of their industry $\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right)$, and how the earnings of their industry differ from the grand mean of earnings

[11] For each 7-year interval, we create percentiles {1,2,…,99} for $\left(y_t^{i,j,k} - \bar{y}_t\right)$, where percentile "X" is defined as the mean of $\left(y_t^{i,j,k} - \bar{y}_t\right)$ for all individuals between the "X - ½" and the "X + ½" percentiles.

$\left(\bar{y}_t^k - \bar{y}_t\right)$. We do this for each percentile in the first and third intervals, and then calculate the difference between the first and third intervals for each percentile.

For each percentile, the black line in Figure 2 is the person component $\Delta\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right)$, the red line is the firm component $\Delta\left(\bar{y}_t^{j,k} - \bar{y}_t^k\right)$, and the blue line is the industry component $\Delta\left(\bar{y}_t^k - \bar{y}_t\right)$. We see that at the lower end of the earnings distribution, industry accounts for most of the declining earnings. At the higher end of the earnings distribution, industry also plays a sizeable role in accounting for increasing earnings. Looking ahead to the detailed analysis of industry, Figure 2 suggests that industry plays a major role in understanding earnings change at both the lower and the upper ends of the earnings distribution.

Of interest is the role of the between-firm, within-industry component in Figure 2. This firm component $\Delta\left(\bar{y}_t^{j,k} - \bar{y}_t^k\right)$ has only a modest contribution to the changing earnings distribution for the first 85 percentiles. The absolute value of the red line is less than .025 for each of the first 87 percentiles. From the 88th to the 99th percentiles, the' firm component increases monotonically to a value of 0.107 for the final percentile.[12]

III. Estimating AKM Fixed Effects Regressions

To further understand the role of workers and firms in the generation of earnings inequality, we rely on the linear model of Abowd, Kramarz, and Margolis (1999). We estimate our model separately for each of three seven-year periods: 1996-2002, 2004-2010, and 2012-2018. Following Song et al. (2019), we assume that earnings $y_{i,j,t}$ are the sum of the effect $\theta_{i,p}$ of worker i in period p, a firm effect $\psi_{j,p}$ when employed by j in period p, and a vector of time-

[12] The analogous percentile figures for males and females are in Appendix A and B respectively.

varying observable characteristics $X_{i,t}$ for worker i at time t, which have distinct marginal effects β_p by period p. We can express this as

$$y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p + \varepsilon_{i,j,t}.$$

Our observable characteristics control for time and worker age. Specifically, we include a set of year dummies that capture calendar year effects on earnings. To control for worker age, we follow the specification of Card, Cardoso, and Kline (2016). We center age around 40, include a quadratic and cubic transformation of worker age, but omit the linear term. To solve this model, we implement the iterative method proposed by Guimaraes and Portugal (2010).

Following the recent literature (e.g., Card, Cardoso and Kline (2016) and Song et. al. (2019)) we estimated the AKM decomposition separately for males and females. Consistent with this literature, all of the AKM based results that follow are reported separately for males and females. We find that qualitatively and quantitatively the results are similar for males and females. To facilitate comparisons with Song et. al. (2019) who focus on results for males in the main text of their paper and report results for females in an appendix, we do the same.[13]

IV. Person, Firm, and Covariance Effects

Table 2 exploits the AKM decomposition of earnings to decompose rising earnings inequality using the person, firm, and covariance effects among males. The first three columns of the table show results for alternative subperiods while the last column computes the terms

[13] Results in Sections IV, V, and VI present results for males. AKM results for females are in Appendix B. Since the results are so similar, we plan on estimating and reporting results in the next draft using AKM results pooling males and females. The pooled results are in progress.

underlying the change in inequality from our first to last subperiods (1996-02 to 2012-18). Rising between-firm dispersion dominates the rise in overall earnings inequality.

The terms in the between-firm dispersion are interpretable as reflecting sorting, segregation, and firm premia effects. Sorting reflects an increased covariance between person and firm effects, given by $2cov(\bar{\theta}^j, \Psi^j) + 2cov(\Psi^j, \overline{X\beta}^j)$. Segregation reflects increased concentration of workers of the same type (captured by person effects), given by $var(\bar{\theta}^j) + var(\overline{X\beta}^j) + 2cov(\bar{\theta}^j, \overline{X\beta}^j)$. The remaining contributor to between-firm dispersion is changing dispersion in firm premia given by $var(\Psi^j)$.

Estimates of the sorting, segregation, and firm premia effects are given in Table 3. Looking first at variance growth from the mid-to-late 1990s to the most recent period, segregation contributes 37.4%, sorting contributes 35.3%, and the rising firm premia contributes 11.8%.

These segregation, sorting, and firm premia results are very similar to Song et. al. (2019) for similar time periods. Using the Song et. al. (2019) results for the subperiods that most closely overlap with ours (1994-2000 to 2007-13), they find that 86.5 percent of variance growth for males is between firms, which is very similar to our result of 84.5 percent. Estimates of the sorting, segregation, and firm premia effects are also similar between Song et al. (2019) and our results for similar time periods. For example, for the variance growth from the mid-to-late 1990s to the most recent period, segregation contributes 35.5% in the Song et.al analysis and 37.3% in our LEHD data. Sorting contributes 37.5% in the Song et.al analysis and 35.3% in our LEHD data. Rising dispersion in firm premium contributes 14.6% in Song et.al and 11.8% in the LEHD. These contributions are broadly similar to those in the longer time interval also reported in Song et. al. from the 1980s to the most recent period with one notable exception. They find a smaller role for firm premia for the longer time interval.

V. Adding Industry to the Variance Decomposition

We now turn to decomposing the contribution of the between-firm components into between-industry and between-firm, within-industry components. Table A.1 in Appendix A presents all the specific components. Of greater interest and more readily interpretable are the between-industry and between-firm, within-industry contributions of sorting, segregation, and firm premia. These results are presented in Table 4.

We already know from Table 1 that between-firm dispersion accounts for 84.5% of the increase in overall earnings inequality and between-industry dispersion accounts for 65.6 percentage points of the 84.5 percentage points. Table 4 shows that 43.3% (28.4 percentage points of the 65.6 percentage points) of this is accounted for by increased sorting of high wage workers into industries with high average firm premia (and sorting of low wage workers into industries with low average firm premia), 42.2% is accounted for increased segregation of workers by person effects across industries, and 14.3% is accounted for by increased dispersion in industry-level average firm premia across industries.

Table 4 also shows that increased sorting and segregation account for most of the 18.9 percentage point increase in between-firm, within-industry dispersion. For the remainder of the main text of the paper, we focus on understanding the dominant role of between-industry effects. Appendix A includes further analysis of the within-industry, between-firm component of rising earnings dispersion.

VI. Understanding the Dominant Role of Industry

To shed further light on the dominant role of industry, we examine the specific industries that contribute the most to the rising between-industry dispersion. Using the earlier notation, between-industry variance growth is $\Delta var\left(\bar{y}^k - \bar{y}\right)$, which is empirically estimated as:[14]

$$\sum_{k=1}^{301} \Delta[\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2].$$

We define industry k's contribution to between-industry variance growth as $\Delta[\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2]$.

Tables 5 and 6 provide statistics and decompositions for the top ten industries contributing to the rising between-industry dispersion, along with the top three industries with the largest (in absolute value) drag on increasing earnings dispersion. Columns 1 and 2 in Table 5 show how the industry's earnings differ from the grand mean $\left(\bar{y}_t^k - \bar{y}_t\right)$, expressed as an average of the first and third intervals and the change from the first to the third interval. Columns 3 and 4 show the average and change of industry k's employment share (N_k/N). Columns 5 and 6 show the variance contribution as a level and as a percent of total between-industry variance growth (0.0829),

The top ten contributing industries account for 61% of the increase in between-industry dispersion while only accounting for 19% of total employment.[15] Six of these ten industries are high-earnings industries with increasing earnings, and four are low-earnings industries with flat or declining earnings or zero earnings change. The high-earnings industries in the top ten for males are Other Information Services, Software Publishers, Computer Systems Design,

[14] The Δ operator is the difference between third and first subperiods. Time subscripts are omitted for convenience.
[15] The calculations of employment shares for the top industries are based on the employment share averaged over the first and third sub-periods.

Management of Companies, Other Financial Investment Activity, and Scientific Research Services. The low-earnings industries in the top ten for males are Restaurants and Other Eating Places, Employment Services, Other General Merchandise Stores, and Grocery Stores. Worth noting is the Restaurants and Other Eating Places industry, which accounts for 18.8 percent of total between-industry variance growth. This 18.8 percent is more than twice as large as the next largest industry contribution. In Appendix B, we show that there is considerable overlap in the top ten contributing industries for females.

Sorting and segregation play important roles in all the top ten contributing industries (Table 6). The contribution of sorting ranges from 33.3% in Computer Systems Design to 49.7% in Grocery Stores. Variation in the contribution of segregation ranges from 27.0% in Other Information Services to 62.3% in Scientific Research Services. Dispersion in industry-level average firm pay premia range from 1.3% in Scientific Research Services to 24.1% in Other Information Services.

In interpreting these findings, it is important to emphasize that the sorting, segregation and firm premia in Table 6 reflect between-industry variation. For example, in the Restaurants and Other Eating Places industry, increased sorting accounts for 48% of the large (18.8%) contribution of this industry to rising between industry dispersion. This reflects an increase in the employment-share weighted cross-product of the average deviation of the firm and person effects in the industry. Relatedly, a substantial segregation contribution for an industry implies a substantial increase in the employment-share weighted squared deviation of the average person effect. A substantial firm premia contribution implies a substantial increase in the employment-share weighted squared deviation of the average firm effect. Given these relationships, it is not surprising that sorting makes the largest contribution in industries with both large segregation and firm premium effects.[16]

Table 7 presents the contribution of all 301 industries by the range of their contributions. Seven industries contribute more than five percent each with a cumulative contribution of 59 percent. Twenty-one industries contribute between one percent and five percent each with a cumulative contribution of almost 40 percent. Thus, the top 28 industries (about ten percent of all industries) account for almost 100 percent of the increase in between-industry dispersion. The top 28 industries account for only 32 percent of total employment.

About 90 percent of the 4-digit NAICS industries each contribute less than one percent in absolute value. This includes about 49 percent of the 4-digit industries each of which contribute zero percent,[17] about 23 percent that each contribute between zero percent and one percent, and about 19 percent that each contribute less than zero percent. The contributions to rising between-industry variance from the latter two groups cancel each other out.

The industries in Table 7 with a negative contribution are only a modest drag on rising earnings inequality. Still the industries with the most negative contribution (see Table 5) are an interesting group. Two of the three most negative are high-earnings high-tech manufacturing industries (Navigational Instruments Manufacturing and Computer Manufacturing) that have had a substantial contraction in employment. The other, Investigation and Security Services, is a low-earnings industry with a substantial increase in earnings.

[16] For the top 28 industries that contribute virtually all of the between industry contribution, the correlation between the segregation and sorting contribution is 0.92, the correlation between the firm premia and sorting contribution is 0.76, and the correlation between the segregation and firm premium contributions is 0.57.

[17] These 148 industries classified as zero percent each contribute less than .00000415 in absolute value to the growth in between-industry variance growth (.00000415/.0829 = .00005, which is the threshold between 0.0 percent and 0.1 percent when rounding to one significant digit after the decimal point).

We look at the 28 industries who each contribute more than 1% of between-industry variance growth by whether they are high-paying or low-paying industries. In Table 8, the 19 high-paying industries contribute 51 percent of between-industry variance growth, and the 9 low-paying industries contribute 47 percent of between-industry variance growth. Amongst these top 28 industries that account for virtually all of the rising between-industry inequality, there are systematic patterns in the contributions of earnings changes vs. employment changes between the 19 high-paying and 9 low-paying industries. Using a shift-share decomposition, we highlight these systematic patterns in Table 8.[18] For the 19 high-paying industries, earnings changes dominate. In contrast, for the 9 low-paying industries, employment changes dominate. Segregation effects are more important in the high-paying industries, sorting is more important in the low-paying industries, and increased dispersion in firm pay premia is more important in the low-paying industries.

The contributions of these 19 high-paying and 9 low-paying industries to earnings dispersion is explored further in Table 9. For comparison, we also include the other 273 industries. The dominant industries had substantial increases in their employment share: the high-paying industries increased their share of employment by 3.31 percentage points, while low-paying industries increased their share by 4.72 percentage points. The other industries decreased their employment share. Of particular note are the other high-paying industries, where the employment share declined by 6.30 percentage points. The dominant industries also became more extreme in terms of their position in the earnings distribution: the average earnings of the

[18] Industry k's contribution to between-industry variance growth is $\Delta\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2$. The shift share term for industry k's changing employment share is $\left[\overline{\left(\bar{y}^k - \bar{y}\right)^2} * \Delta\left(\frac{N_k}{N}\right) / \Delta\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2\right]$. The shift share term for industry k's changing earnings is $\left[\overline{\left(\frac{N_k}{N}\right)} * \Delta\left(\bar{y}^k - \bar{y}\right)^2 / \Delta\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2\right]$.

19 high-paying industries increased by 0.1596, while that of the low-paying industries declined by 0.1143. Overall, the 19 high-paying industries account for 51.4% of the between-industry increase in earnings dispersion, while the 9 low-paying industries account for nearly all of the remainder: 47.2%.

To sum up, the top ten percent of contributing industries account for virtually all of the increase in between-industry dispersion. The top ten percent tend to be in the tails of the earnings distribution. The top contributing industries with high earnings exhibit especially increasing earnings differentials. The top contributing industries with low earnings exhibit especially increasing employment. Increased sorting and segregation accounts for most of the increased between-industry dispersion in the top contributing industries, but rising dispersion in average industry-level firm premia also plays an important supporting role. For the dominant low-paying industries, increasing sorting is relatively more important than increased segregation.

VII. The Role of Mega Firms[19]

Much attention has been given to the increased share of economic activity accounted for by the largest firms in the economy (see, e.g., Autor et. al. 2020). In this section, we show that changing employment share and changing earnings size premia especially for the mega (10,000+) firms plays a critical role in accounting for rising between-industry earnings inequality.

[19] Song et. al. (2019) also examine the role of mega firms but with a different focus. They do not explore the close connection between rising between industry dispersion and mega firms in a relatively narrow set of industries. They do note however that rising within firm inequality is greater at the mega firms. We find in Appendix A that the industries that contribute most to rising between industry dispersion also contribute the most to rising between firm, within industry inequality. We have not yet quantified the role of mega firms for this pattern but the patterns we report in this section suggest this channel is important. We plan to investigate this more directly in the next draft.

Figure 3 reports the changes in employment share by size class for the four groups of industries in Table 8. Strikingly, the rising employment share of mega firms is dominated by the relatively small share of industries that account for virtually all of the increase in between-industry earnings inequality. For the 19 high-paying industries in the top ten percent of contributing industries, the employment share of mega firms has increased by 4.5 percentage points from the 1996-2002 to 2012-18 periods. For the 9 low-paying industries in the top ten percent of contributing industries, the employment share has increased by 8.4 percentage points. In contrast, the remaining 271 industries have exhibited modest declines in the share of employment in the mega firms: the employment share of mega firms increased by 0.7 percentage points for the 147 other low-paying industries, and decreased by 1.9 percentage points for the 126 other high-paying industries.

The 19 top paying industries exhibits a rising size-earnings premium (Figure 4). The size premium rises for all size classes between our first subperiod (1996-2002) and our third subperiod (2012-2018). For the 19 top paying industries, earnings rise by 16 to 17 percent by size classes 250-999, by 19 percent for size class 1000-9999, and by 16 percent for mega firms. These increase are due to both increases in the AKM firm and person effects. The 9 low-paying industries exhibit a decline in the size-earnings premium, with declines of 16 percent at mega firms. Both AKM firm and person effects contribute to the declining premium for these 9 low-paying industries.

For the 147 other low-paying industries, there is also a decrease in the size-earnings premium accounted for primarily by a decrease in the person effect. The magnitude of the decline in the size-earnings premium is much smaller for these industries, with estimates relatively close to zero for the small and medium size classes and then declining to a 9 percent loss for the mega firms. For the 126 other high-paying industries, there is not a monotonic change in the size-earnings premium. These industries have an average earnings gain of 5 percent.

Figure 5 provides the underpinnings of Figure 4 for the low-paying and high paying industries that dominate rising between industry inequality. At the dominant low paying industries, firm premia are negative (below economy wide average) as are the average person effect. The firm premium and average person effect were less negative at mega firms than at other firms in the 1990s, but this has become less true over time yielding a flattening of the relationship between earnings and size. In the late 1990s, the firm premium for mega firms in these low pay industries exceeded that at the smallest size class (20-99) by 5 percentage points. By 2012-16, this gap had narrowed to 2 percentage points. Relatedly, the average person effect at the mega firms was 20 percentage higher than at the smallest size class in the 1990s. This gap narrowed to 11 percentage points by 2012-18. It is also notable that the average person effect declined substantially for all size classes over this period of time. These patterns highlight that a core contributing factor to rising earnings inequality is that a relatively small number (9) of low paying industries became even lower paying especially at the mega firms. This decline in the size-earnings premium is accompanied by a sharp increase in employment in the mega firms.

Turning to the two lower panels of Figure 5, mega firms had a firm premium about 11 percentage points higher than the smallest size class for the top paying industries that dominate rising between industry earnings inequality in the late 1990s. This premium increased to 16 percentage points by 2012-18. Mega firms had about 9 percentage points higher person fixed effect than the smallest size class in the 1990s and this gap increased to 11 percentage points by 2012-18. These patterns highlight that for the small number (19) of top paying industries that dominate rising between industry dispersion the mega firms increased their relative firm premium and average person effect substantially relative to smaller firms in the same industry.

To put these results into perspective with the findings of a declining overall size-earnings premium reported in Bloom et. al. (2018), Figure 6 reports the earnings changes and employment share changes for all industries pooled together. We also find a rising share of employment at mega firms (2.5 percentage points) accompanied by an inverted U-shaped change in the size-earnings premium, with particularly large declines for mega firms. However, pooling across all industries masks several key results that are evident in Figures 4 and 5. First, the increase in employment share at mega firms is concentrated in the ten percent of industries that account for rising between-industry earnings inequality. Second, the 19 high-paying industries exhibit a pronounced increase in the size-earnings premium rather than a decrease. For these industries, the interaction of a rising employment share with an increase in the size-earnings premium helps account for the rising earnings differentials exhibited by these industries. Third, the 9 low-paying industries have the largest increase in the employment share of mega firms. This increase combined with the falling size-earnings premium helps account for the falling earnings differentials exhibited by these industries.

VIII. Concluding Remarks

Rising earnings inequality is dominated by rising between-firm inequality. Our analysis as well as the recent literature emphasizes that this largely reflects how firms are organizing themselves in terms of their workforce. High (low) earnings workers are more likely to work with each other (increased segregation), and high (low) earnings workers are more likely to work at high (low) firm premia firms (sorting).

Our contribution is to highlight the dominant role of industry effects in accounting for this structural change of how firms organize their workforces.[20] Most of the rising between-firm inequality is accounted for by rising between-industry dispersion in earnings. The between-industry component accounts for two thirds of increasing earnings inequality, and 78 percent of between-firm inequality growth. This changes the narrative of the sorting and segregation contributions. High (low) earnings workers are more likely to work with each other in specific industries and high (low) earnings workers are more likely to work in high (low) average firm premia industries.

Not only do industry effects dominate but it is a relatively small share of industries that account for virtually all the increasing dispersion in earnings across industries. We find that about ten percent of the 301 detailed 4-digit NAICS industries account for about 100% of the rising between-industry dispersion, while accounting for less than 40% of employment. The ten percent of industries that account for virtually all of the increase are drawn from the top and bottom of the earnings distribution in terms of industry-level averages. For those industries at the top of the earnings distribution, their contribution is dominated by rising inter-industry earnings differentials. For industries at the bottom of the earnings distribution, their contribution is dominated by shifts in employment to these very low-earnings industries. For both sets of industries at the top and the bottom of the earnings distribution, increased sorting and segregation between industries dominates but increased dispersion in between-industry firm premia also plays an important supporting role. Increased sorting is relatively more important for the rising between industry dispersion from the industries at the bottom of the earnings

[20] The summary of results in the concluding remarks focuses on the results reported in the main text for males. Results for females are reported in appendix B and are very similar.

distribution. In contrast, increased segregation is relatively more important for the rising between industry dispersion from the industries at the top of the earnings distribution.

The dominance of industry effects is closely linked to the rising importance of mega (10,000+) firms in the U.S. economy. The increasing share of employment accounted for by mega firms is concentrated in the 28 industries that account for virtually all of rising between-industry dispersion. This rising share of employment at mega firms is accompanied by a declining size-earnings premium in the 9 low-paying industries in the top 28 but a rising size-earnings premium in the 19 high-paying industries in the top 28.

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. Moreover, since it is the between industry contribution that dominates, it is the common effects of re-organization across firms in the same industry that matter. Many mechanisms such as changing technology, market structure, and globalization likely underlie rising earnings inequality. The focus of future research on the impact of such changes on rising earnings inequality should be on the uneven and concentrated impact of such mechanisms across industries. Moreover, the focus of such research should be on mechanisms that can also account for the rising share of activity of mega firms and the changing size-earnings premium in this relatively small number of industries. For the top paying dominant industries, mega firms are gaining market share and increasing the quality of their workers as well as increasing the firm premium relative to smaller firms in these industries. For the low-paying dominant industries, mega firms are also gaining market share but are decreasing the quality of their workers as well as decreasing the firm premium relative to smaller firms in these industries. In short, our analysis implies that understanding the role of mega firms in a small number of industries is also critical for understanding the sources of rising overall earnings inequality in the economy.

These mechanisms mostly operate via increased sorting and segregation between industries in the relatively small set of industries that account for most of the rising between-industry dispersion. However, rising dispersion in between-industry firm premia also plays an important supporting role, especially in the high-paying industries. Importantly, the increased sorting and segregation reflect between-industry effects that are concentrated in a relatively small number of industries.

Our findings imply that the role of inter-industry earnings differentials and the changing composition of employment across industries is much more important for understanding earnings inequality than suggested by the recent literature. Our findings are derived from comprehensive matched employer-employee administrative data with high quality industry codes from Census and BLS processing of employer data. Our results contrast not only with the recent studies using SSA data but also with findings from household surveys such as the Current Population Survey (CPS). Stansbury and Summers (2020) find that there is rising between-industry earnings dispersion in the raw CPS but find this is reversed when they control for individual and occupation characteristics. Hoffmann, Lee, and Lemieux (2020) present closely related evidence from the CPS showing declining between-industry earnings dispersion after controlling for individual and occupation characteristics.[21]

[21] Haltiwanger and Spletzer (2020a) find rising residual between-industry dispersion using a two-step procedure that first controls for observable worker and firm characteristics and then controls for occupation effects using the OES data. The AKM approach used in the current paper dominates this two step procedure but the two-step procedure uses the same high quality industry codes as in the current paper. Moreover, the two-step procedure is closer methodologically to the approaches in the recent literature using the CPS.

Although analysis using the CPS data controls for individual and occupation characteristics, we note that our analysis is based on the AKM firm-level premium component, which abstracts from person effects as well as the covariance between the firm-level and person-level effects. Further investigation of this stark difference in findings from the CPS and our findings using administrative matched employer-employee data is needed, ideally using matched CPS-LEHD microdata. We note, however, that it has long been recognized that there are systematic limitations of industry codes in household surveys (see Mellow and Sider (1983) and Dey et. al. (2010)).[22] Our findings highlight the importance of using high quality, detailed industry codes for drawing inferences about the changing structure of earnings and employment across industries.

[22] We also note that a major difference between earnings data from the CPS versus from administrative data is trouble in the tails (Bollinger et.al, 2019). Specifically, nonresponse in the CPS is U-shaped, with low-earnings persons and high-earnings persons being the least likely to report their earnings. Bollinger et.al. find that this nonresponse can account for up to one-half of the differences in measured inequality between the CPS and administrative SSA data. An investigation of the role of industry in accounting for increasing inequality would need to pay particular attention to industries at the tails of the earnings distribution.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. "High Wage Workers and High Wage Firms." *Econometrica*, Vol. 67, No. 2, pp, 251-333.

Abowd, John M., Kevin L. McKinney, and Nellie L. Zhao. 2018. "Earnings Inequality and Mobility Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data." *Journal of Labor Economics*, Vol. 36, No. 51, pp. S183-S300.

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock. 2009. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators," in *Producer Dynamics*, Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts, eds., Chicago: University of Chicago Press, pp. 149-230.

Autor David, David Dorn, Lawrence F. Katz, Christina C. Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *Quarterly Journal of Economics*, Vol. 135, No. 2, pp. 645-709.

Barth, Erling, Alex Bryson, James C, Davis, and Richard Freeman. 2016. "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States." *Journal of Labor Economics* 34 (2, pt. 2): pp. S67-S97.

Bloom, Nicholas, Fatih Guvenen, Benjamin S. Smith, Jae Song, and Till von Wachter. 2018. "The Disappearing Large-Firm Premium." *American Economic Review Papers and Proceedings*, 108: pp. 317-322.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2019. "Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch." *Journal of Political Economy*, Vol. 127, No. 5, pp. 2143-2185.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics*, Vol. 131, No. 2, pp. 633-686.

Card, David, Jorg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics*, Vol. 128, No. 3, pp. 967-1015.

Davis, Steve J., and John Haltiwanger. 1991. "Wage dispersion between and within U.S. manufacturing plants, 1963–1986." *Brookings Papers on Economic Activity: Microeconomics*, pp. 115–200.

Dey, Matthew, Susan Houseman, and Anne Polivka. 2010. "What Do We Know About Contracting Out in the United States? Evidence from Household and Establishment Surveys" in *Labor in the New Economy,* Katharine G. Abraham, James R. Spletzer, and Michael Harper, eds., Chicago: University of Chicago Press, pp. 267-304.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Ken Troske. 2004. "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment." *Journal of Labor Economics*, pp. 397-429.

Guimaraes, Paulo and Pedro Portugal. 2010. "A simple feasible procedure to fit models with high-dimensional fixed effects." *The Stata Journal*, Vol. 10, No. 4, pp. 628-649.

Haltiwanger, John and James Spletzer. 2020a. "Between Firm Changes in Earnings Inequality: The Dominant Role of Industry Effects." *NBER Working Paper No.* 26786, February.

Haltiwanger, John and James Spletzer. 2020b. "Rising Between Firm Inequality and Declining Labor Market Fluidity: Evidence of a Changing Job Ladder." Paper presented at NBER/CRIW conference on Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth, March.

Hoffman, Florian, David S. Lee, and Thomas Lemieux. 2020. "Growing Income Inequality in the United States and Other Advanced Economies." *Journal of Economic Perspectives*, Vol. 34, No. 4, pp. 52-78.

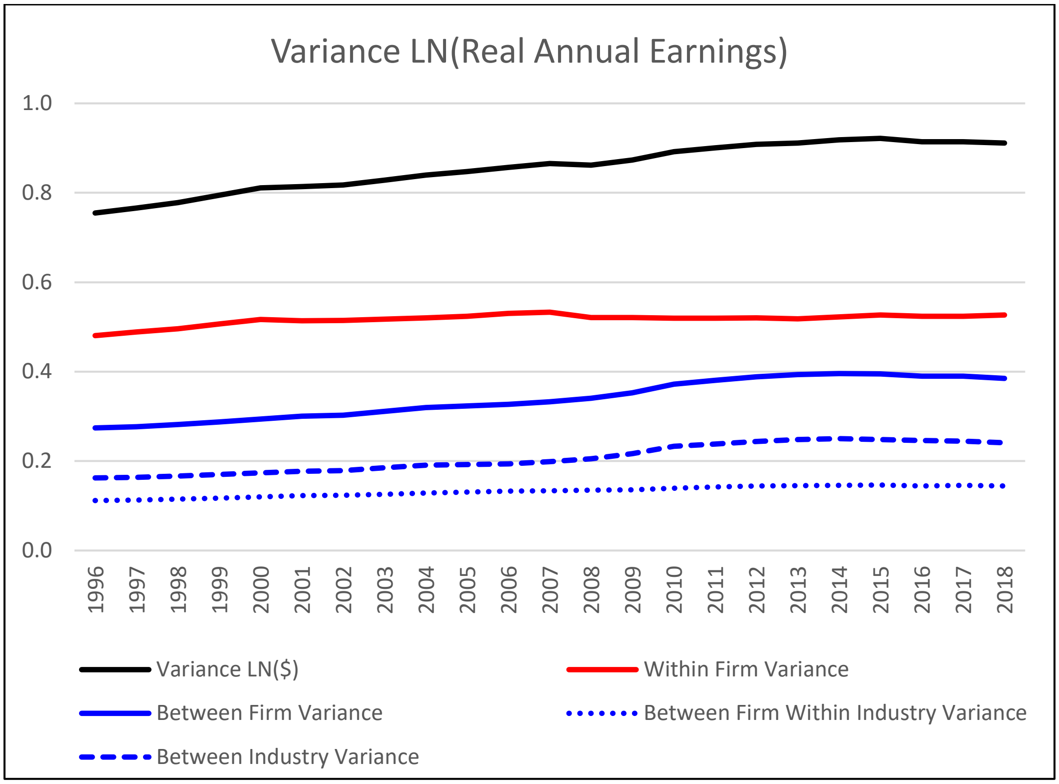
Mellow, Wesley, and Hal Sider. 1983. "Accuracy of response in labor market surveys: Evidence and implications." *Journal of Labor Economics* 1, no. 4:331–44.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2019. "Firming Up Inequality." *Quarterly Journal of Economics*, Vol. 134, No. 1, pp. 1–50.

Stansbury, Anna, and Lawrence H. Summers. 2020. "The Declining Worker Power Hypothesis: An explanation for the recent evolution of the American economy." *Brookings Papers on Economic Activity*, Spring, pp. 1-77.

Walker, Ed. 1997. "The Census Bureau's Business Register: Basic Features and Quality Issues." Paper presented at the Joint Statistical Meetings of the American Statistical Association, Anaheim, CA.

Figure 1: Variance Decomposition by Year, 1996-2018 (All Workers)



Notes: (1) Persons with annual real earnings >$3770 in EINs with 20 or more employees. 2003 and 2011 are linearly interpolated.

(2) $var\left(y_t^{i,j,k} - \bar{y}_t\right) = var\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right) + var\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right) + var\left(\bar{y}_t^{k} - \bar{y}_t\right)$

Table 1: Variance Decomposition by 7-Year Interval (All Workers)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j,k})$	.7938	.8623	.9152	.1214
Person $var(y^{i,j,k} - \bar{y}^{j,k})$	.5124	.5319	.5305	.0181
Firm $var(\bar{y}^{j,k} - \bar{y}^{k})$	.1115	.1271	.1396	.0281
Industry $var(\bar{y}^{k} - \bar{y}_t)$	.1699	.2033	.2451	.0752
Person $var(y^{i,j,k} - \bar{y}^{j,k})$	64.6%	61.7%	58.0%	14.9%
Firm $var(\bar{y}^{j,k} - \bar{y}^{k})$	14.0%	14.7%	15.3%	23.1%
Industry $var(\bar{y}^{k} - \bar{y}_t)$	21.4%	23.6%	26.8%	61.9%
Sample size (millions)	239.4	249.2	269.7	
Number of firms (thousands)	470	460	466	
Number of NAICS industries	301	301	301	

Notes: (1) Persons with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $var(y_t^{i,j,k} - \bar{y}_t) = var(y_t^{i,j,k} - \bar{y}_t^{j,k}) + var(\bar{y}_t^{j,k} - \bar{y}_t^{k}) + var(\bar{y}_t^{k} - \bar{y}_t)$

Table 1A: Variance Decomposition by 7-Year Interval (Males)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var\left(y^{i,j,k}\right)$	.8358	.9107	.9621	.1263
Person $var\left(y^{i,j,k} - \bar{y}^{j,k}\right)$	.5299	.5527	.5495	.0196
Firm $var\left(\bar{y}^{j,k} - \bar{y}^{k}\right)$	.1277	.1441	.1516	.0239
Industry $var\left(\bar{y}^{k} - \bar{y}_t\right)$	.1782	.2139	.2610	.0828
Person $var\left(y^{i,j,k} - \bar{y}^{j,k}\right)$	63.4%	60.7%	57.1%	15.5%
Firm $var\left(\bar{y}^{j,k} - \bar{y}^{k}\right)$	15.3%	15.8%	15.8%	18.9%
Industry $var\left(\bar{y}^{k} - \bar{y}_t\right)$	21.3%	23.5%	27.1%	65.6%
Sample size (millions)	131.7	135.0	145.7	
Number of firms (thousands)	470	460	466	
Number of NAICS industries	301	301	301	

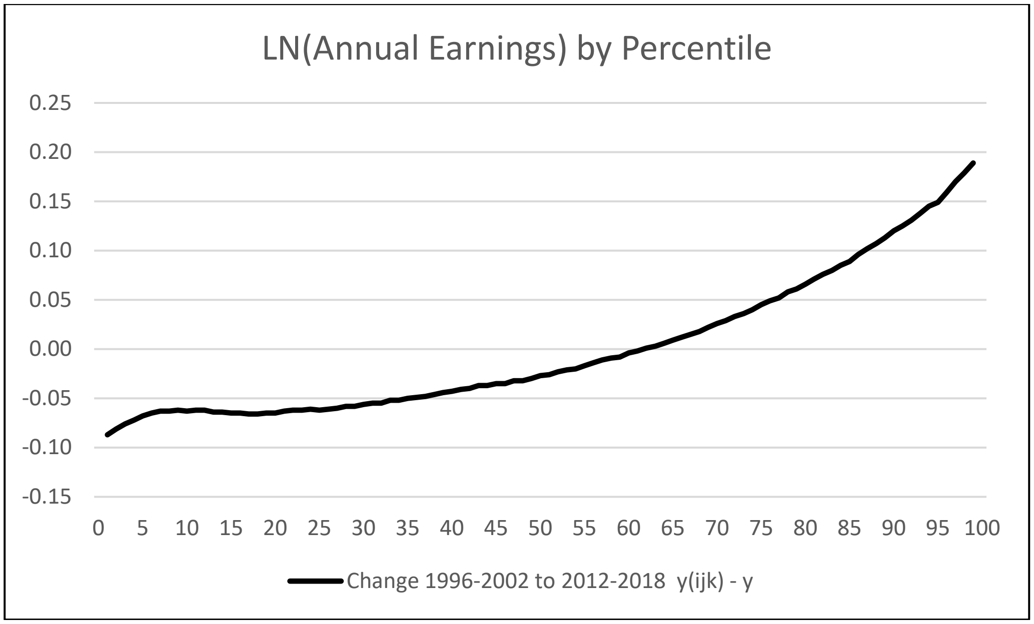
Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $var\left(y_t^{i,j,k} - \bar{y}_t\right) = var\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right) + var\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right) + var\left(\bar{y}_t^{k} - \bar{y}_t\right)$

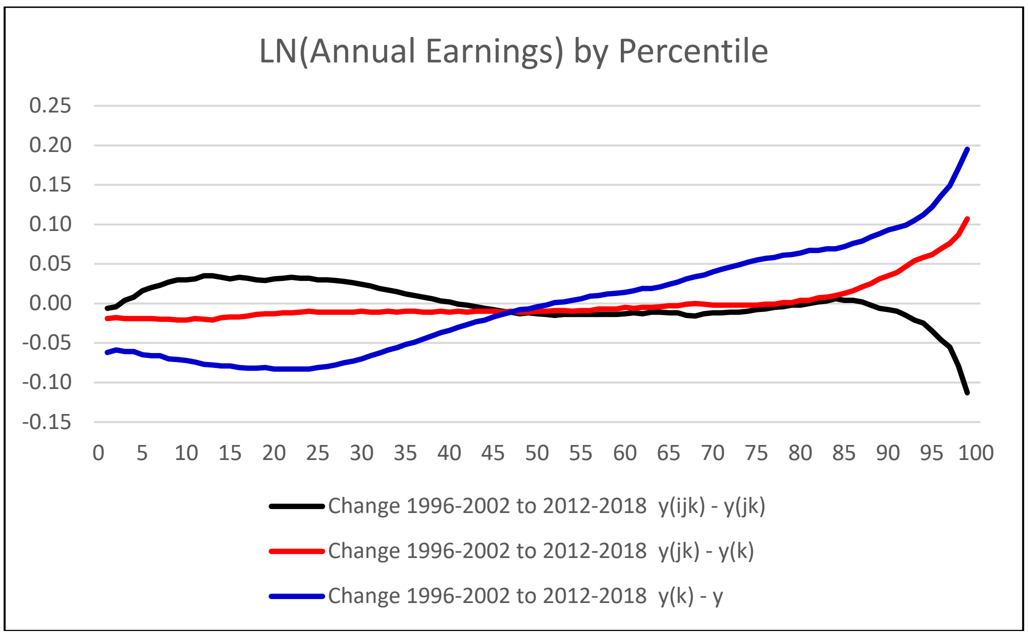
Table 1B: Variance Decomposition by 7-Year Interval (Females)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j,k})$	.6676	.7463	.8068	.1392
Person $var(y^{i,j,k} - \bar{y}^{j,k})$	.4338	.4629	.4736	.0398
Firm $var(\bar{y}^{j,k} - \bar{y}^{k})$	.0944	.1104	.1267	.0323
Industry $var(\bar{y}^{k} - \bar{y}_{t})$	.1394	.1730	.2065	.0671
Person $var(y^{i,j,k} - \bar{y}^{j,k})$	65.0%	62.0%	58.7%	28.6%
Firm $var(\bar{y}^{j,k} - \bar{y}^{k})$	14.1%	14.8%	15.7%	23.2%
Industry $var(\bar{y}^{k} - \bar{y}_{t})$	20.9%	23.2%	25.6%	48.2%
Sample size (millions)	107.7	114.2	124.0	
Number of firms (thousands)	470	460	466	
Number of NAICS industries	301	301	301	

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $var(y_t^{i,j,k} - \bar{y}_t) = var(y_t^{i,j,k} - \bar{y}_t^{j,k}) + var(\bar{y}_t^{j,k} - \bar{y}_t^{k}) + var(\bar{y}_t^{k} - \bar{y}_t)$

Figure 2: Change LN(Real Annual Earnings) by Percentile (All Workers)





Notes: (1) Persons with annual real earnings >$3770 in EINs with 20 or more employees.

(2) $\left(y_t^{i,j,k} - \bar{y}_t\right) = \left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right) + \left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right) + \left(\bar{y}_t^{k} - \bar{y}_t\right)$

(3) For each 7-year interval, we create percentiles {1, 2, …, 99} for $\left(y_t^{i,j,k} - \bar{y}_t\right)$, where percentile "X" is defined as the mean of $\left(y_t^{i,j,k} - \bar{y}_t\right)$ for all individuals between the "X - ½th" and the "X + ½th" percentiles. For each of these 99 percentiles, we create the means of $\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right)$, $\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right)$, and $\left(\bar{y}_t^{k} - \bar{y}_t\right)$.

Table 2: Variance Decomposition (Males)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.8358	.9107	.9621	.1263
Between-firm $var(\bar{y}^{j})$	36.6%	39.3%	42.9%	84.5%
$var(\bar{\theta}^{j})$	11.5%	12.2%	13.2%	23.9%
$var(\Psi^{j})$	7.4%	8.3%	8.0%	11.8%
$var(\overline{X\beta}^{j})$	1.2%	1.1%	1.3%	1.8%
$2cov(\bar{\theta}^{j}, \Psi^{j})$	10.8%	11.4%	12.9%	26.6%
$2cov(\bar{\theta}^{j}, \overline{X\beta}^{j})$	2.7%	3.0%	3.8%	11.6%
$2cov(\Psi^{j}, \overline{X\beta}^{j})$	3.0%	3.3%	3.7%	8.7%
Within-firm $var(y^{i,j} - \bar{y}^{j})$	63.4%	60.7%	57.1%	15.5%
$var(\theta^{i} - \bar{\theta}^{j})$	40.9%	39.2%	37.3%	13.6%
$var(X\beta^{i,j} - \overline{X\beta}^{j})$	8.5%	6.5%	7.8%	3.0%
$var(\epsilon^{i,j})$	15.7%	14.7%	13.5%	-1.0%
$2cov(\theta^{i} - \bar{\theta}^{j}, X\beta^{i,j} - \overline{X\beta}^{j})$	-1.9%	0.1%	-1.4%	2.0%
$2cov(\theta^{i} - \bar{\theta}^{j}, \epsilon^{i,j})$	0.2%	0.2%	-0.1%	-2.0%
$2cov(X\beta^{i,j} - \overline{X\beta}^{j}, \epsilon^{i,j})$	0.1%	0.1%	0.1%	-0.2%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p$

Table 3: Variance Decomposition (Males)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.8358	.9107	.9621	.1263
Between-firm $var(\bar{y}^{j})$	36.6%	39.3%	42.9%	84.5%
Firm segregation	15.4%	16.3%	18.3%	37.4%
Firm pay premium	7.4%	8.3%	8.0%	11.8%
Firm sorting	13.8%	14.8%	16.6%	35.3%
Within-firm $var(y^{i,j} - \bar{y}^{j})$	63.4%	60.7%	57.1%	15.5%
Person effect	47.5%	45.8%	43.7%	18.6%
Residual	15.9%	14.9%	13.4%	-3.1%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p$
(3) Firm segregation defined as $var(\bar{\theta}^{j})+var(\overline{X\beta}^{j})+2cov(\bar{\theta}^{j}, \overline{X\beta}^{j})$.
Firm pay premium defined as $var(\Psi^{j})$.
Firm sorting defined as $2cov(\bar{\theta}^{j}, \Psi^{j,k}) + 2cov(\Psi^{j}, \overline{X\beta}^{j,k})$.
Person effect defined as $var(\theta^{i} - \bar{\theta}^{j})+ var(X\beta^{i,j} - \overline{X\beta}^{j})+$
$2cov(\theta^{i} - \bar{\theta}^{j}, X\beta^{i,j} - \overline{X\beta}^{j})$.
Residual defined as $var(\epsilon^{i,j,k})+2cov(\theta^{i} - \bar{\theta}^{j}, \epsilon^{i,j})+$
$2cov(X\beta^{i,j} - \overline{X\beta}^{j}, \epsilon^{i,j})$.

Table 4: Industry Enhanced Variance Decomposition (Males)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j,k})$	.8358	.9107	.9621	.1263
Between-firm, within-industry $var(\bar{y}^{j,k} - \bar{y}^{k})$	15.3%	15.8%	15.8%	18.9%
Firm segregation	7.6%	7.8%	7.9%	9.7%
Firm pay premium	3.4%	3.7%	3.3%	2.4%
Firm sorting	4.2%	4.3%	4.6%	6.9%
Between-industry $var(\bar{y}^{k})$	21.3%	23.5%	27.1%	65.6%
Industry segregation	7.8%	8.4%	10.4%	27.7%
Industry pay premium	4.0%	4.6%	4.7%	9.4%
Industry sorting	9.5%	10.5%	12.0%	28.4%
Within-firm $var(y^{i,j,k} - \bar{y}^{j,k})$	63.4%	60.7%	57.1%	15.5%
Person effect	47.5%	45.8%	43.7%	18.6%
Residual	15.9%	14.9%	13.4%	-3.1%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

(2) $y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p$

(3) Firm segregation defined as $var(\bar{\theta}^{j,k} - \bar{\theta}^{k})$+$var(\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$+ $2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}), (\overline{X\beta}^{j,k} - \overline{X\beta}^{k})]$.

Firm pay premium defined as $var(\Psi^{j,k} - \bar{\Psi}^{k})$.

Firm sorting defined as $2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}), (\Psi^{j,k} - \bar{\Psi}^{k})]$+ $2cov[(\Psi^{j,k} - \bar{\Psi}^{k}),](\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$.

Industry segregation defined as $var(\bar{\theta}^{k})$+$var(\overline{X\beta}^{k})$+$2cov(\bar{\theta}^{k}, \overline{X\beta}^{k})$.

Industry pay premium defined as $var(\bar{\Psi}^{k})$.

Industry sorting defined as $2cov(\bar{\theta}^{k}, \bar{\Psi}^{k})$+$2cov(\bar{\Psi}^{k}, \overline{X\beta}^{k})$.

Person effect defined as $var(\theta^{i} - \bar{\theta}^{j,k})$+ $var(X\beta^{i,j,k} - \overline{X\beta}^{j,k})$+ $2cov(\theta^{i} - \bar{\theta}^{j,k}, X\beta^{i,j,k} - \overline{X\beta}^{j,k})$.

Residual defined as $var(\epsilon^{i,j,k})$+$2cov(\theta^{i} - \bar{\theta}^{j,k}, \epsilon^{i,j,k})$+ $2cov(X\beta^{i,j,k} - \overline{X\beta}^{j,k}, \epsilon^{i,j,k})$.

Table 5: Industry Contributions to Between-Industry Variance Growth, Top 10 and Bottom 3 Industries (Males)

NAICS	Industry Title	Average $(\bar{y}^k - \bar{y})$	Change $(\bar{y}^k - \bar{y})$	Average (N_k/N)	Change (N_k/N)	Variance Contribut	Var Cont % of Total
7225	Restaurants & Other Eating Places	-.8038	-.0577	.0436	.0179	.0156	18.8%
5613	Employment Services	-.8137	.0000	.0374	.0105	.0069	8.4%
4529	Other General Merchandise Stores	-.5681	-.1358	.0112	.0130	.0060	7.2%
5191	Other Information Services	.8112	.7077	.0023	.0034	.0053	6.4%
5112	Software Publishers	.9662	.1737	.0063	.0032	.0051	6.2%
5415	Computer Systems Design	.6111	.0200	.0216	.0121	.0050	6.1%
4451	Grocery Stores	-.3903	-.2709	.0220	.0010	.0048	5.8%
5511	Management of Companies	.4787	.1837	.0205	-.0006	.0035	4.2%
5239	Other Financial Investment Activity	.9091	.3418	.0025	.0017	.0030	3.7%
5417	Scientific Research Services	.7394	.2348	.0082	-.0007	.0024	2.9%
3345	Navigational Instruments Manuf	.6529	.0578	.0092	-.0038	-.0009	-1.1%
5616	Investigation and Security Services	-.5669	.1374	.0104	.0021	-.0009	-1.1%
3341	Computer Manufacturing	.8653	.1698	.0063	-.0040	-.0011	-1.4%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

(2) $(\bar{y}^k - \bar{y})$ is the industry's quarterly LN(real annual earnings) average relative to the economy average.

(3) (N_k/N) is the industry's employment share.

(4) Industry k's contribution to between-industry variance growth is $\Delta[\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2]$, which sums over industries to .0829.

Table 6: Industry Contributions to Between-Industry Variance Growth, Top 10 and Bottom 3 Industries (Males)

NAICS	Industry Title	Variance Contribut	Segregat	Pay Premium	Sorting	Shift Share Earnings	Shift Share Emp Share
7225	Restaurants & Other Eating Places	18.8%	38.3%	13.7%	48.0%	25.9%	74.1%
5613	Employment Services	8.4%	40.8%	13.2%	46.0%	0.0%	100.0%
4529	Other General Merchandise Stores	7.2%	36.2%	16.3%	47.5%	28.9%	71.1%
5191	Other Information Services	6.4%	27.0%	24.1%	48.9%	50.2%	49.8%
5112	Software Publishers	6.2%	43.5%	11.6%	44.9%	41.1%	58.9%
5415	Computer Systems Design	6.1%	63.2%	3.5%	33.3%	10.5%	89.5%
4451	Grocery Stores	5.8%	29.8%	20.6%	49.7%	96.4%	3.6%
5511	Management of Companies	4.2%	51.0%	7.3%	41.7%	104.3%	-4.3%
5239	Other Financial Investment Activity	3.7%	48.2%	9.5%	42.3%	51.2%	48.8%
5417	Scientific Research Services	2.9%	62.3%	1.3%	36.5%	116.3%	-16.3%
3345	Navigational Instruments Manuf	-1.1%	14.9%	29.6%	55.4%	-73.9%	173.9%
5616	Investigation and Security Services	-1.1%	3.2%	22.9%	74.0%	171.7%	-71.7%
3341	Computer Manufacturing	-1.4%	8.8%	37.0%	54.3%	-161.5%	261.5%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

(2) Industry k's contribution to between-industry variance growth is $\Delta\left(\frac{N_k}{N}\right)[(\bar{y}^k-\bar{y})^2]$, which sums over industries to .0829.

(3) The shift share term for changing employment share is $\left[\overline{(\bar{y}^k-\bar{y})^2} * \Delta\left(\frac{N_k}{N}\right)/\Delta[\left(\frac{N_k}{N}\right)(\bar{y}^k-\bar{y})^2]\right]$. The shift share term for changing earnings is $\left[\overline{\left(\frac{N_k}{N}\right)} * \Delta(\bar{y}^k-\bar{y})^2/\Delta[\left(\frac{N_k}{N}\right)(\bar{y}^k-\bar{y})^2]\right]$.

Table 7: Industry Contributions to Between-Industry Variance Growth, By Variance Contribution (Males)

Contribution to Variance		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
>5%	7 industries	.0489	59.0%	39.5%	14.5%	46.1%	32.2%	67.8%
1% - 5%	21 industries	.0328	39.6%	41.9%	16.1%	42.0%	92.2%	7.8%
0% - 1%	69 industries	.0169	20.4%	37.5%	-2.7%	45.0%	91.1%	8.9%
=0%	148 industries	-.0001	-0.2%					
<0%	56 industries	-.0156	-18.8%	28.3%	22.6%	49.1%	11.4%	88.6%
		.0829	100.0%	42.3%	14.4%	43.3%	73.5%	26.5%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Cells in the fourth row are blank due to dividing by a something close to zero.

Table 8: Industry Contributions to Between-Industry Variance Growth, By Variance Contribution and Average Earnings (Males)

		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
28 industries with variance contribution >1%								
High-paying	19 industries	.0426	51.4%	44.5%	14.2%	41.2%	73.6%	26.4%
Low-Paying	9 industries	.0391	47.2%	36.0%	16.1%	47.9%	37.5%	62.5%
273 industries with variance contribution ≤1%								
High-paying	126 industries	.0008	0.9%					
Low-Paying	147 industries	.0004	0.5%					
		.0829	100.0%	42.3%	14.4%	43.3%	73.5%	26.5%

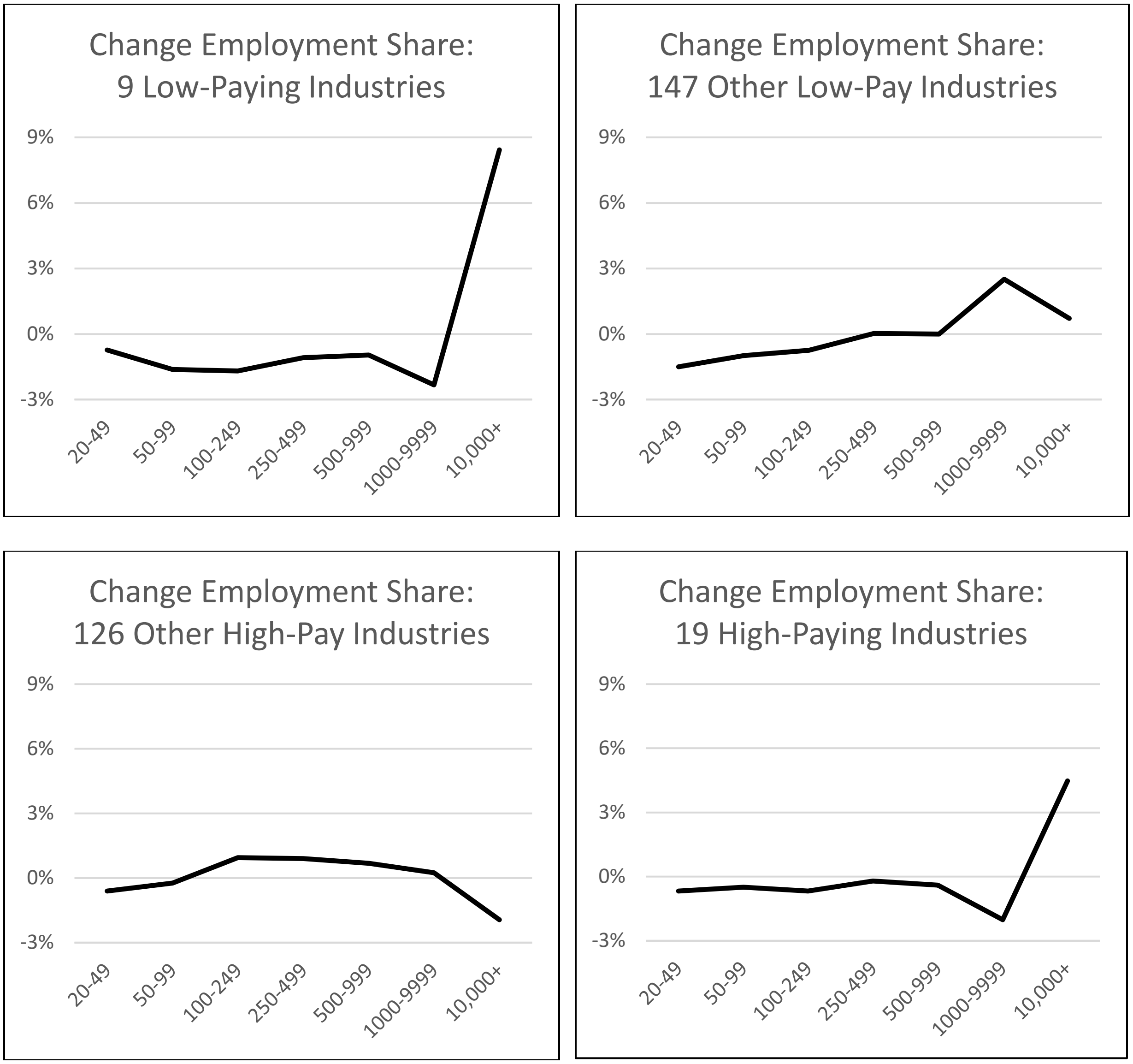
Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Cells in the last two rows are blank due to dividing by a something close to zero.

Table 9: Industry Contributions to Between-Industry Variance Growth, By Variance Contribution and Average Earnings (Males)

	Average $(\bar{y}^k - \bar{y})$	Change $(\bar{y}^k - \bar{y})$	Average (N_k/N)	Change (N_k/N)	Variance Contribut	Var Cont % of Total
28 industries with variance contribution >1%						
High Paying 19 industries	.5317	.1596	.1762	.0331	.0426	51.4%
Low-Paying 9 industries	-.6643	-.1143	.1418	.0472	.0391	47.2%
273 industries with variance contribution ≤1%						
High Paying 126 industries	.2725	.0414	.3493	-.0630	.0008	0.9%
Low-Paying 147 industries	-.2825	-.0009	.3329	-.0168	.0004	0.5%
					.0829	100.0%

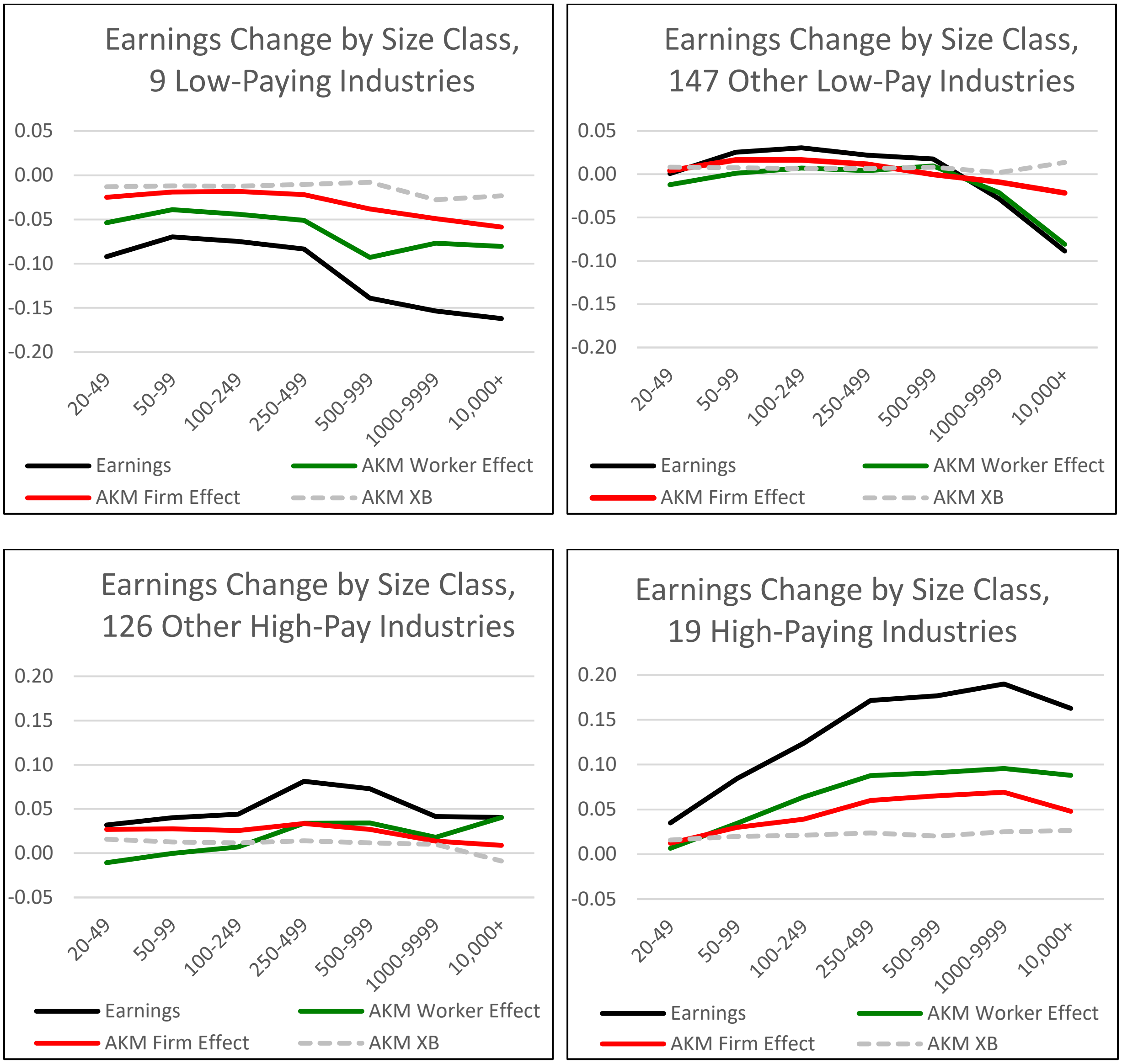
Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

Figure 3: Changing Employment Share by Size Class for Select Industries (Males)



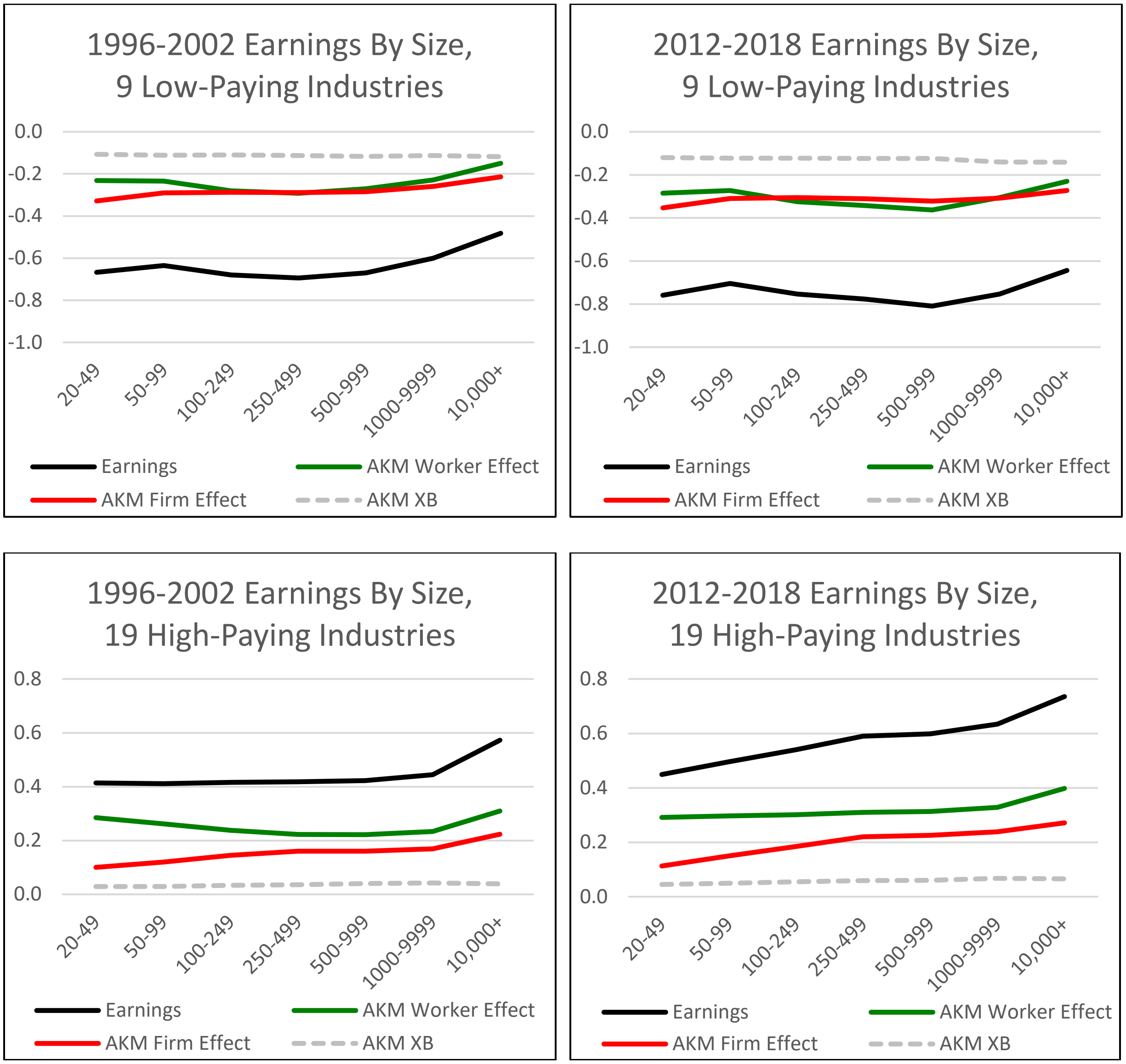
Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

Figure 4: Earnings Change by Size Class for Select Industries (Males)



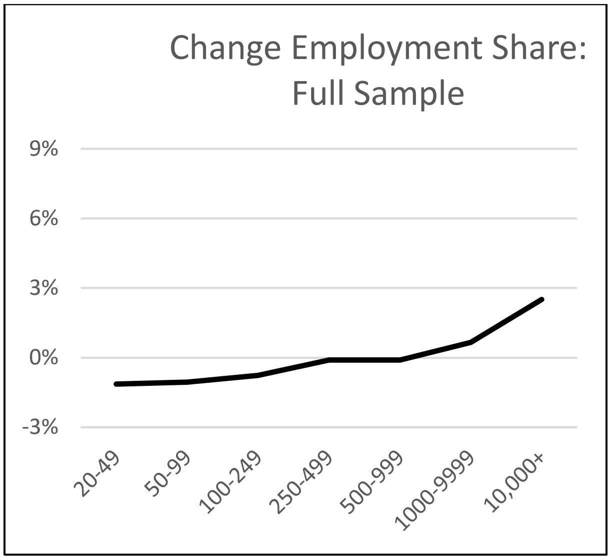
Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

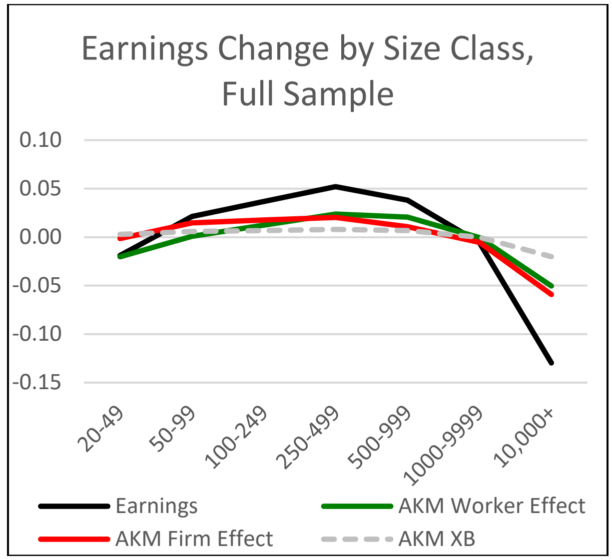
Figure 5: Earnings Levels by Size Class for Select Industries (Males)



Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

Figure 6: Changing Employment Share and Earnings Change by Size Class for Full Sample of All Industries (Males)





Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

Appendix A

This appendix includes supplemental tables and figures for the results highlighted in the main text (Figure A.1 and Table A.1) as well as discussion of the between firm, within industry component of rising earnings inequality. The discussion in this appendix focuses on the latter. Given the main text includes results for males, the results in this section use males results only. Results for females are in Appendix B.

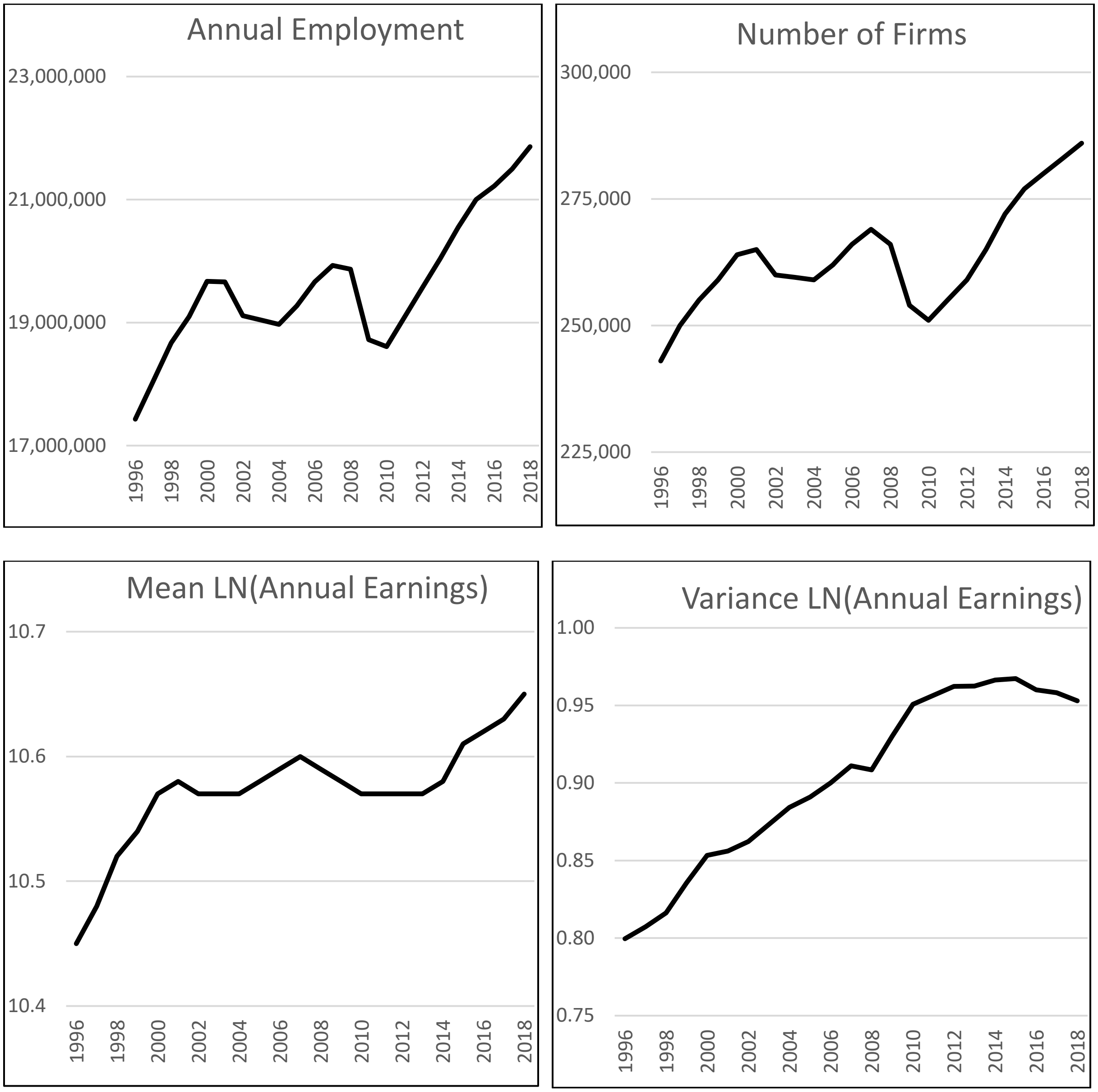
Table A.2 and A.3 are the analogs of Tables 5 and 6 but for the between firm, within industry contribution. The top ten industries alone contribute 65% to the between firm, within industry component while accounting for only 17% of employment. Four of the top ten industries in Table A.2 are also among the top ten industries (for the between industry component) in Table 5. These industries include Employment Services, Computer Systems Design, Restaurants and Other Eating Places, and Other Information Services. For the six non-overlapping 4-digit industries, three overlap at the 3-digit or 2-digit level.

The overlap in the ranking of industries in terms of the between industry component and between firm, within industry component is far from perfect. A good example of this is Investigation and Security Services which is in the bottom three for the between industry component (contributing negatively) and in the top ten for the between firm, within industry component. This is a low-earnings industry that has exhibited a substantial increase in average earnings (see Table 5). However, within the industry, there has been a shift in employment towards the higher earnings firms within the industry along with rising dispersion of between firm earnings differentials within the industry.

While there is a strong relationship between the magnitude of the between firm, within industry components and the between industry components, the between industry components are much smaller in magnitude. This translates into a slope coefficient in Figure A.2 of 0.16.

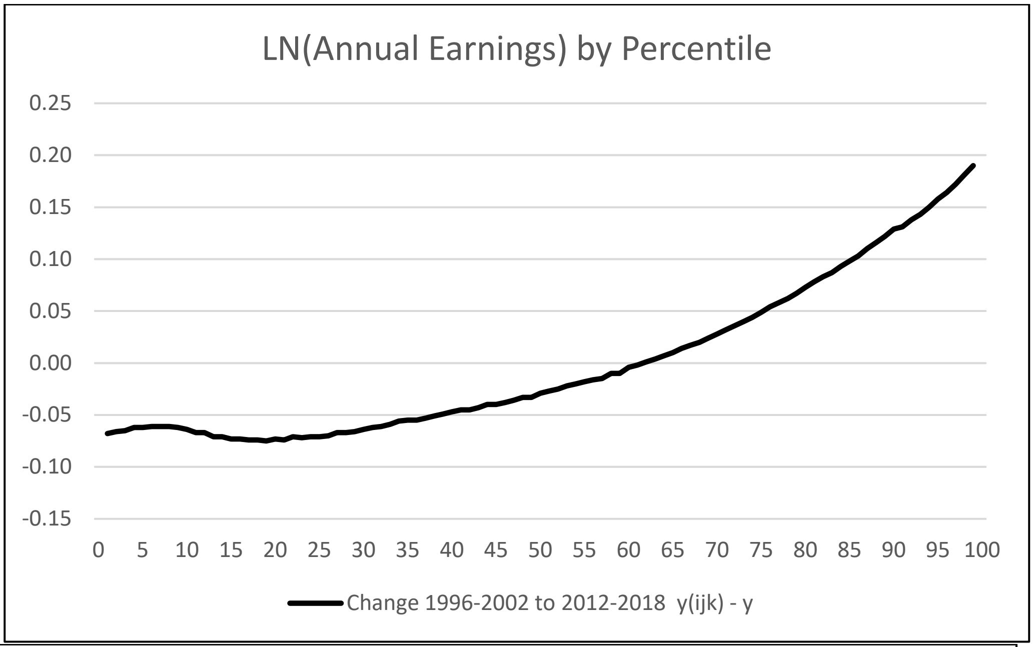
Tables A.4 and A.5 illustrate that the within industry, between firm component is also concentrated in a relatively small fraction of industries. The top 36 industries with a contribution in excess of 1% account for more than 100% of the overall within industry, between firm contribution. 24 of the top 36 are high-paying industries, and in contrast to the between industry high-paying industry results, employment changes are relatively more important than earnings changes in accounting for their contribution. For the 12 low paying industries in the top 36, employment changes are also relatively more important than earnings changes.

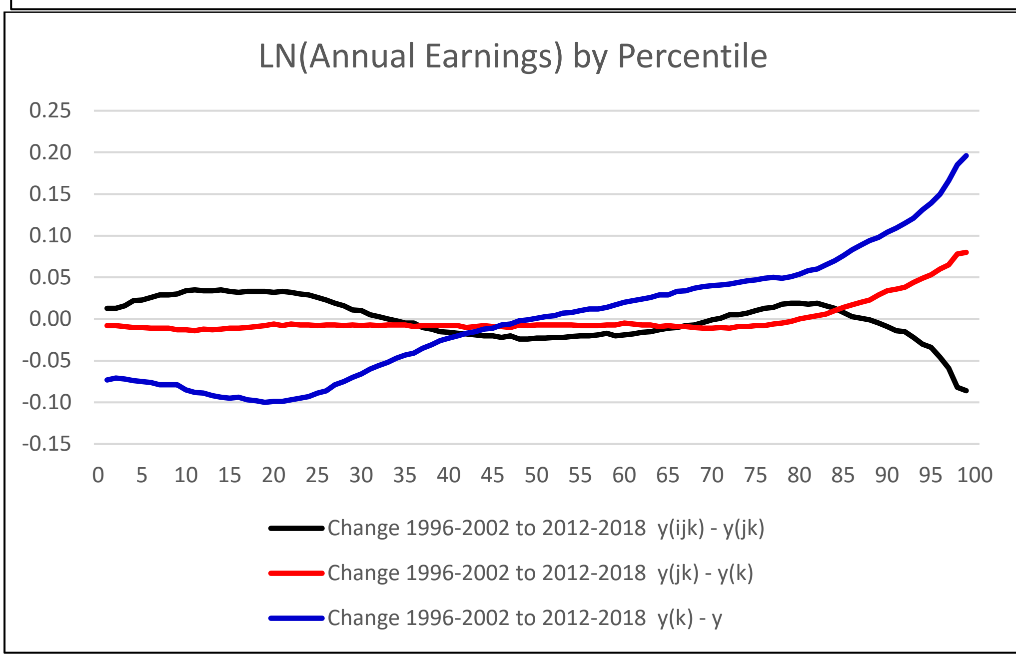
Figure A.1: Descriptive Statistics (Males)



Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

Figure A.2: Change LN(Real Annual Earnings) by Percentile (Males)





Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

(2) $\left(y_t^{i,j,k} - \bar{y}_t\right) = \left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right) + \left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right) + \left(\bar{y}_t^{k} - \bar{y}_t\right)$

(3) For each 7-year interval, we create percentiles $\{1, 2, \ldots, 99\}$ for $\left(y_t^{i,j,k} - \bar{y}_t\right)$, where percentile "X" is defined as the mean of $\left(y_t^{i,j,k} - \bar{y}_t\right)$ for all individuals between the "X - ½th" and the "X + ½th" percentiles. For each of these 99 percentiles, we create the means of $\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right)$, $\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right)$, and $\left(\bar{y}_t^{k} - \bar{y}_t\right)$.

Table A.1: Industry Enhanced Variance Decomposition (Males)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j,k})$	.8358	.9107	.9621	.1263
Between-firm, within-industry $var(\bar{y}^{j,k} - \bar{y}^{k})$	15.3%	15.8%	15.8%	18.9%
$var(\bar{\theta}^{j,k} - \bar{\theta}^{k})$	6.0%	6.2%	6.2%	7.6%
$var(\Psi^{j,k} - \bar{\Psi}^{k})$	3.4%	3.7%	3.3%	2.4%
$var(\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$	0.7%	0.5%	0.6%	0.1%
$2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}),(\Psi^{j,k} - \bar{\Psi}^{k})]$	3.4%	3.4%	3.7%	5.9%
$2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}),(\overline{X\beta}^{j,k} - \overline{X\beta}^{k})]$	0.9%	1.1%	1.1%	2.0%
$2cov[(\Psi^{j,k} - \bar{\Psi}^{k}),](\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$	0.8%	0.9%	0.9%	1.0%
Between industry $var(\bar{y}^{k})$	21.3%	23.5%	27.1%	65.6%
$var(\bar{\theta}^{k})$	5.6%	5.9%	7.0%	16.3%
$var(\bar{\Psi}^{k})$	4.0%	4.6%	4.7%	9.4%
$var(\overline{X\beta}^{k})$	0.5%	0.5%	0.7%	1.7%
$2cov(\bar{\theta}^{k}, \bar{\Psi}^{k})$	7.4%	8.0%	9.1%	20.7%
$2cov(\bar{\theta}^{k}, \overline{X\beta}^{k})$	1.7%	1.9%	2.8%	9.7%
$2cov(\bar{\Psi}^{k}, \overline{X\beta}^{k})$	2.2%	2.5%	2.9%	7.7%
Within firm $var(y^{i,j,k} - \bar{y}^{j,k})$	63.4%	60.7%	57.1%	15.5%
$var(\theta^{i} - \bar{\theta}^{j,k})$	40.9%	39.2%	37.3%	13.6%
$var(X\beta^{i,j,k} - \overline{X\beta}^{j,k})$	8.5%	6.5%	7.8%	3.0%
$var(\epsilon^{i,j,k})$	15.7%	14.7%	13.5%	-1.0%
$2cov(\theta^{i} - \bar{\theta}^{j,k}, X\beta^{i,j,k} - \overline{X\beta}^{j,k})$	-1.9%	0.1%	-1.4%	2.0%
$2cov(\theta^{i} - \bar{\theta}^{j,k}, \epsilon^{i,j,k})$	0.2%	0.2%	-0.1%	-2.0%
$2cov(X\beta^{i,j,k} - \overline{X\beta}^{j,k}, \epsilon^{i,j,k})$	0.1%	0.1%	0.1%	-0.2%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p$

Table A.2: Industry Contributions to Within-Industry Variance Growth, Top 10 and Bottom 3 Industries (Males)

NAICS	Industry Title	Average $(\bar{y}^k - \bar{y})$	Change $(\bar{y}^k - \bar{y})$	Average (N_k/N)	Change (N_k/N)	Variance Contribut	Var Cont % of Total
5613	Employment Services	-.8137	.0000	.0374	.0105	.0028	11.7%
5415	Computer Systems Design	.6111	.0200	.0216	.0121	.0024	10.2%
5416	Management Scientific … Services	.3871	.0543	.0085	.0062	.0021	8.8%
7225	Restaurants & Other Eating Places	-.8038	-.0577	.0436	.0179	.0018	7.7%
5191	Other Information Services	.8112	.7077	.0023	.0034	.0012	5.2%
4541	Electronic Shopping & Mail-Order	.1407	.4631	.0023	.0016	.0011	4.7%
5121	Motion Picture & Video Industries	-.0230	-.0921	.0048	.0009	.0010	4.3%
5616	Investigation & Security Services	-.5669	.1374	.0104	.0021	.0010	4.3%
6211	Offices of Physicians	.7111	.0190	.0078	.0024	.0010	4.3%
4251	Wholesale Electronic Markets	.2915	.1549	.0039	.0011	.0009	3.9%
3231	Printing & Support Activities	.0770	-.0976	.0070	-.0051	-.0005	-1.9%
4431	Electronics & Appliance Stores	-.2400	-.1624	.0058	-.0021	-.0006	-2.4%
3341	Computer Manufacturing	.8653	.1698	.0063	-.0040	-.0006	-2.6%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $(\bar{y}^k - \bar{y})$ is the industry's quarterly LN(real annual earnings) average relative to the economy average.
(3) (N_k/N) is the industry's employment share.
(4) Industry k's contribution to within-industry variance growth is $\Delta[\left(\frac{N_k}{N}\right)\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k} - \bar{y}^k\right)^2]$, which sums over industries to .0239.

Table A.3: Industry Contributions to Within-Industry Variance Growth, Top 10 and Bottom 3 Industries (Males)

NAICS	Industry Title	Variance Contribut	Segregat	Pay Premium	Sorting	Shift Share Earnings	Shift Share Emp Share
5613	Employment Services	.0028	51.4%	11.4%	37.1%	7.1%	92.9%
5415	Computer Systems Design	.0024	61.7%	9.9%	28.4%	11.0%	89.0%
5416	Management Scientific … Services	.0021	53.6%	12.4%	34.0%	-4.1%	104.1%
7225	Restaurants & Other Eating Places	.0018	54.1%	16.4%	29.5%	26.4%	73.6%
5191	Other Information Services	.0012	24.4%	36.6%	39.0%	25.5%	74.5%
4541	Electronic Shopping & Mail-Order	.0011	42.3%	14.4%	43.2%	51.1%	48.9%
5121	Motion Picture & Video Industries	.0010	44.7%	13.6%	41.7%	63.0%	37.0%
5616	Investigation & Security Services	.0010	46.6%	16.5%	36.9%	67.1%	32.9%
6211	Offices of Physicians	.0010	51.0%	4.9%	44.1%	-13.2%	113.2%
4251	Wholesale Electronic Markets	.0009	48.4%	12.9%	38.7%	62.0%	38.0%
3231	Printing & Support Activities	-.0005	48.9%	22.2%	28.9%	7.4%	92.6%
4431	Electronics & Appliance Stores	-.0006	27.6%	24.1%	48.3%	7.9%	92.1%
3341	Computer Manufacturing	-.0006	37.7%	24.6%	37.7%	18.8%	81.2%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

(2) Industry k's contribution to within-industry variance growth is $\Delta[\left(\frac{N_k}{N}\right)\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k}-\bar{y}^{k}\right)^2]$, which sums over industries to .0239.

(3) The shift share term for changing employment share is $\left[\overline{\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k}-\bar{y}^{k}\right)^2} * \Delta\left(\frac{N_k}{N}\right)/\Delta[\left(\frac{N_k}{N}\right)\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k}-\bar{y}^{k}\right)^2]\right]$.

The shift share term for changing earnings is $\left[\overline{\left(\frac{N_k}{N}\right)} * \Delta[\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k}-\bar{y}^{k}\right)^2]/\Delta[\left(\frac{N_k}{N}\right)\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k}-\bar{y}^{k}\right)^2]\right]$.

Table A.4: Industry Contributions to Within-Industry Variance Growth, By Variance Contribution (Males)

Contribution to Variance		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
>5%	5 industries	.0104	43.5%	51.5%	15.1%	33.3%	11.3%	88.7%
1% - 5%	31 industries	.0178	74.5%	45.2%	17.3%	37.5%	42.3%	57.7%
0% - 1%	86 industries	.0080	33.7%	51.6%	16.8%	31.6%	77.5%	22.5%
=0%	68 industries	-.0001	-0.3%					
<0%	111 industries	-.0122	-51.3%	40.3%	24.7%	35.0%	-4.3%	104.3%
		.0239	100.0%	52.1%	12.1%	35.8%	66.4%	33.6%

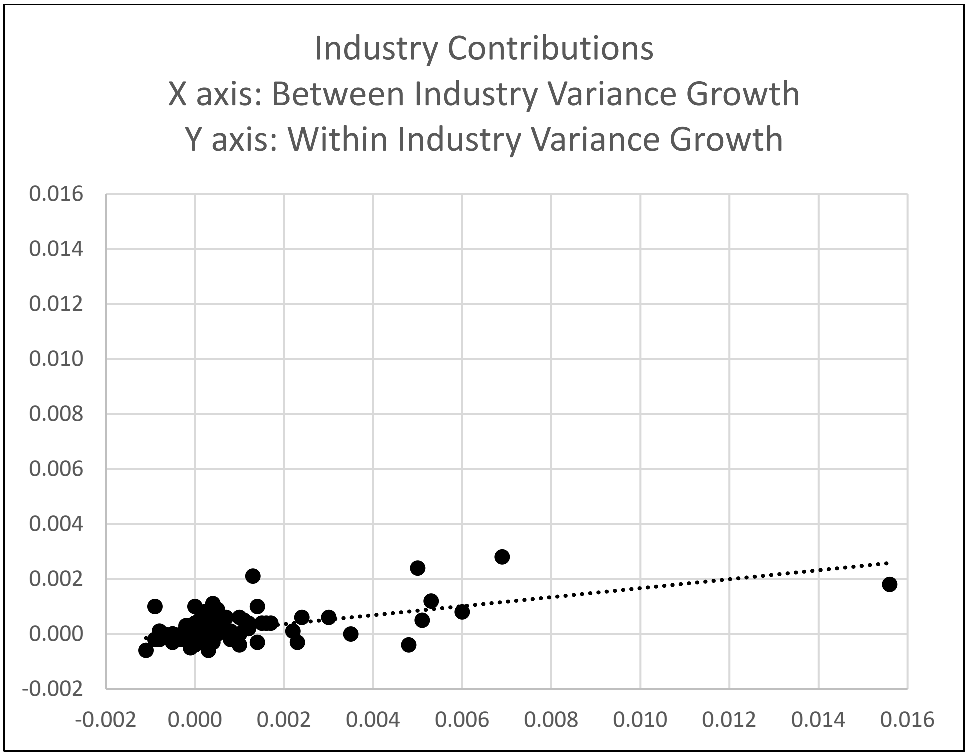
Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Cells in the fourth row are blank due to dividing by a something close to zero.

Table A.5: Industry Contributions to Within-Industry Variance Growth,
By Variance Contribution and Average Earnings (Males)

		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
36 industries with variance contribution >1%								
High Paying	24 industries	.0151	63.3%	48.8%	16.7%	34.5%	25.2%	74.8%
Low-Paying	12 industries	.0130	54.7%	46.1%	16.2%	37.7%	37.5%	62.5%
265 industries with variance contribution ≤1%								
High Paying	141 industries	-.0044	-18.4%					
Low-Paying	124 industries	.0001	0.5%					
		.0239	100.0%	52.1%	12.1%	35.8%	66.4%	33.6%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Cells in the last two rows are blank due to dividing by a something close to zero.

Figure A.2: Industry Contributions to Between-Industry Variance Growth and Within-Industry Variance Growth (Males)



Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.

Appendix B

In this Appendix, we present results for females. To facilitate comparisons with the tables and figures from the main text, the numbering of tables and figures corresponds to tables in the main text. For example, Table B.2 is the version of Table 2 for females. The results for males and females separately are very similar both qualitatively and quantitatively.[23] For both, rising between-industry dispersion is the most important component of rising overall earnings dispersion and the rising between-industry dispersion is concentrated in a relatively small number of industries. The top ten industries accounting for rising between-industry dispersion overlap considerably with seven being the same for both males and females. The exceptions generally show up in the top ten percent of industries that account for virtually all of the increase in between-industry dispersion. The tight relationship between the results for females and males is evident in Figure B.7 depicting a scatterplot of the between-industry contribution for all 301 4-digit industries. The correlation in contribution is 0.76. The estimated slope (standard error) in the scatterplot is 0.56 (0.03).

Rising between-industry dispersion is quantitatively more important for males than females. However, the patterns of the relative contributions of sorting, segregation, and firm premia effects as well as the patterns of contributions in high-paying vs. low-paying industries are very similar.

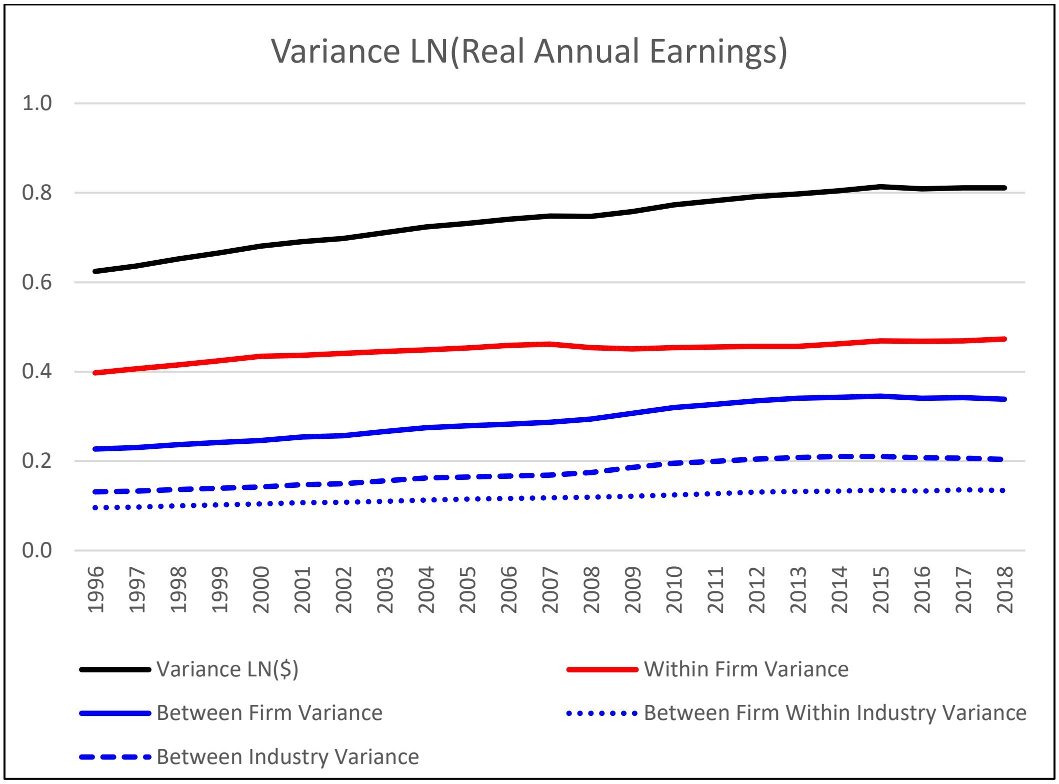
Using the Song et. al. (2019) results for the subperiods that most closely overlap with ours (1994-2000 to 2007-13), they find that 73.4 percent of females earnings variance growth is between firms, which is very similar to our result of 71.4 percent. Estimates of the sorting, segregation, and firm premia effects are also similar between Song et al. (2019) and our results

[23] Results by size class (Figures 3-5 in the main text) have not yet been produced for females. They will be added in a future draft.

for similar time periods. For example, for the variance growth of females from the mid-to-late 1990s to the most recent period, segregation contributes 28.7% in the Song et.al analysis and 31.2% in our LEHD data. Sorting contributes 33.0% in the Song et.al analysis and 32.0% in our LEHD data. Rising dispersion in firm premium contributes 11.7% in Song et.al and 8.3% in the LEHD.

While the findings on the respective contributions of between firm dispersion and the components in terms sorting, segregation, and firm premia match Song et. al. (2019) results closely for females, the key difference is that we find that these patterns reflect between industry effects in a relatively small number of industries. For example, for females as well as males, we find it is increased sorting of industries with low (high) average person effects and low(high) average firm effects in a small number of industries such as restaurants and other eating places (software publishers) that dominates increasing earnings inequality.

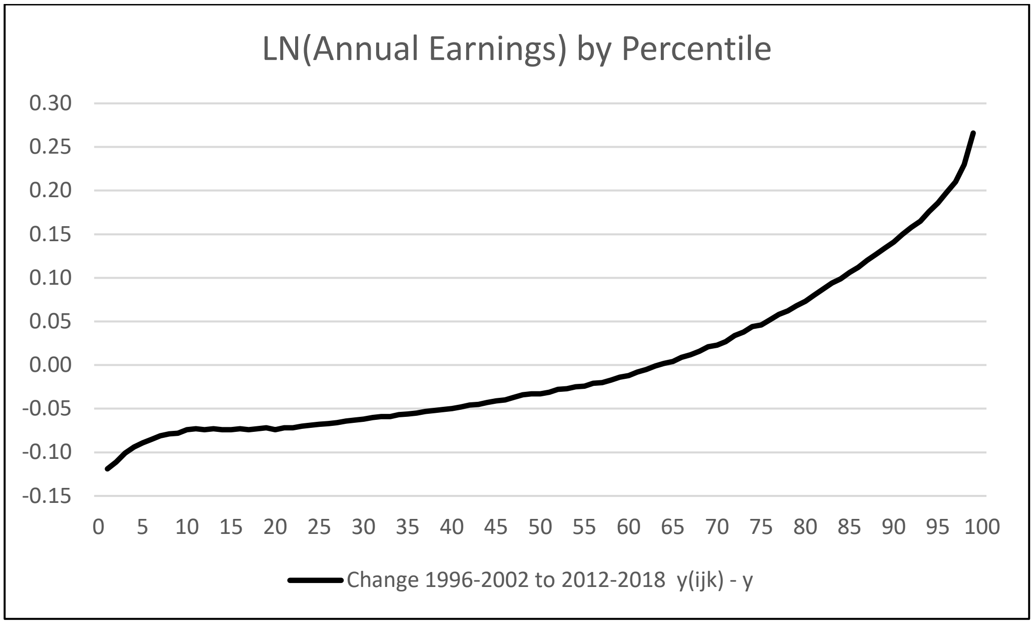
Figure B.1: Variance Decomposition by Year, 1996-2018 (Females)

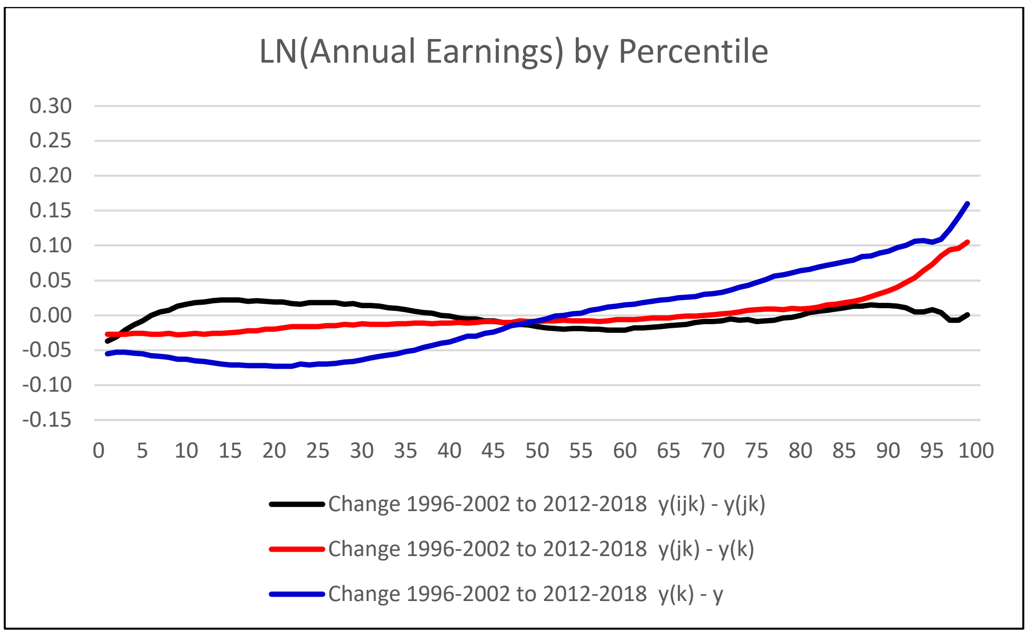


Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees. 2003 and 2011 are linearly interpolated.

(2) $var\left(y_t^{i,j,k} - \bar{y}_t\right) = var\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right) + var\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right) + var\left(\bar{y}_t^{k} - \bar{y}_t\right)$

Figure B.2: Change LN(Real Annual Earnings) by Percentile (Females)





Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

(2) $\left(y_t^{i,j,k} - \bar{y}_t\right) = \left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right) + \left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right) + \left(\bar{y}_t^{k} - \bar{y}_t\right)$

(3) For each 7-year interval, we create percentiles {1, 2, …, 99} for $\left(y_t^{i,j,k} - \bar{y}_t\right)$, where percentile "X" is defined as the mean of $\left(y_t^{i,j,k} - \bar{y}_t\right)$ for all individuals between the "X - ½th" and the "X + ½th" percentiles. For each of these 99 percentiles, we create the means of $\left(y_t^{i,j,k} - \bar{y}_t^{j,k}\right)$, $\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right)$, and $\left(\bar{y}_t^{k} - \bar{y}_t\right)$.

Table B.2: Variance Decomposition (Females)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.6676	.7463	.8068	.1392
Between-firm $var(\bar{y}^j)$	35.0%	38.0%	41.3%	71.4%
$var(\bar{\theta}^j)$	10.1%	10.6%	11.9%	20.3%
$var(\Psi^j)$	8.7%	9.5%	8.7%	8.3%
$var(\overline{X\beta}^j)$	1.1%	0.9%	1.4%	2.9%
$2cov(\bar{\theta}^j, \Psi^j)$	11.1%	12.1%	13.2%	23.1%
$2cov(\bar{\theta}^j, \overline{X\beta}^j)$	1.5%	2.0%	2.6%	8.0%
$2cov(\Psi^j, \overline{X\beta}^j)$	2.5%	3.0%	3.6%	8.9%
Within-firm $var(y^{i,j} - \bar{y}^j)$	65.0%	62.0%	58.7%	28.6%
$var(\theta^i - \bar{\theta}^j)$	41.9%	40.7%	38.8%	23.9%
$var(X\beta^{i,j} - \overline{X\beta}^j)$	7.4%	5.3%	7.1%	6.1%
$var(\epsilon^{i,j})$	18.2%	16.1%	14.6%	-2.7%
$2cov(\theta^i - \bar{\theta}^j, X\beta^{i,j} - \overline{X\beta}^j)$	-2.7%	-0.3%	-2.0%	1.2%
$2cov(\theta^i - \bar{\theta}^j, \epsilon^{i,j})$	0.1%	0.1%	0.1%	0.0%
$2cov(X\beta^{i,j} - \overline{X\beta}^j, \epsilon^{i,j})$	0.1%	0.1%	0.1%	0.1%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p$

Table B.3: Variance Decomposition (Females)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.6676	.7463	.8068	.1392
Between-firm $var(\bar{y}^{j})$	35.0%	38.0%	41.3%	71.4%
Firm segregation	12.7%	13.5%	15.9%	31.1%
Firm pay premium	8.7%	9.5%	8.7%	8.3%
Firm sorting	13.6%	15.0%	16.8%	32.0%
Within-firm $var(y^{i,j} - \bar{y}^{j})$	65.0%	62.0%	58.7%	28.6%
Person effect	46.6%	45.7%	43.9%	31.2%
Residual	18.4%	16.3%	14.8%	-2.6%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

(2) $y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p$

(3) Firm segregation defined as $var(\bar{\theta}^{j})$+$var(\overline{X\beta}^{j})$+$2cov(\bar{\theta}^{j}, \overline{X\beta}^{j})$.

Firm pay premium defined as $var(\Psi^{j})$.

Firm sorting defined as $2cov(\bar{\theta}^{j}, \Psi^{j,k}) + 2cov(\Psi^{j}, \overline{X\beta}^{j,k})$.

Person effect defined as $var(\theta^{i} - \bar{\theta}^{j})$+ $var(X\beta^{i,j} - \overline{X\beta}^{j})$+ $2cov(\theta^{i} - \bar{\theta}^{j}, X\beta^{i,j} - \overline{X\beta}^{j})$.

Residual defined as $var(\epsilon^{i,j,k})$+$2cov(\theta^{i} - \bar{\theta}^{j}, \epsilon^{i,j})$+ $2cov(X\beta^{i,j} - \overline{X\beta}^{j}, \epsilon^{i,j})$.

Table B.4: Industry Enhanced Variance Decomposition (Females)

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var(y^{i,j,k})$	.6676	.7463	.8068	.1392
Between-firm, within-industry $var(\bar{y}^{j,k} - \bar{y}^{k})$	14.1%	14.8%	15.7%	23.2%
Firm segregation	6.2%	6.5%	7.2%	11.8%
Firm pay premium	3.9%	4.1%	3.7%	2.5%
Firm sorting	4.0%	4.2%	4.8%	8.9%
Between-industry $var(\bar{y}^{k})$	20.9%	23.2%	25.6%	48.2%
Industry segregation	6.5%	6.9%	8.7%	19.3%
Industry pay premium	4.8%	5.4%	5.0%	5.8%
Industry sorting	9.6%	10.9%	11.9%	23.1%
Within-firm $var(y^{i,j,k} - \bar{y}^{j,k})$	65.0%	62.0%	58.7%	28.6%
Person effect	46.6%	45.7%	43.9%	31.2%
Residual	18.4%	16.3%	14.8%	-2.6%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

(2) $y_{i,j,t} = \theta_{i,p} + \psi_{j,p} + X_{i,t}\beta_p$

(3) Firm segregation defined as $var(\bar{\theta}^{j,k} - \bar{\theta}^{k})$+$var(\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$+ $2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}), (\overline{X\beta}^{j,k} - \overline{X\beta}^{k})]$.

Firm pay premium defined as $var(\Psi^{j,k} - \bar{\Psi}^{k})$.

Firm sorting defined as $2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}), (\Psi^{j,k} - \bar{\Psi}^{k})]$+ $2cov[(\Psi^{j,k} - \bar{\Psi}^{k}),](\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$.

Industry segregation defined as $var(\bar{\theta}^{k})$+$var(\overline{X\beta}^{k})$+$2cov(\bar{\theta}^{k}, \overline{X\beta}^{k})$.

Industry pay premium defined as $var(\bar{\Psi}^{k})$.

Industry sorting defined as $2cov(\bar{\theta}^{k}, \bar{\Psi}^{k})$+$2cov(\bar{\Psi}^{k}, \overline{X\beta}^{k})$.

Person effect defined as $var(\theta^{i} - \bar{\theta}^{j,k})$+ $var(X\beta^{i,j,k} - \overline{X\beta}^{j,k})$+ $2cov(\theta^{i} - \bar{\theta}^{j,k}, X\beta^{i,j,k} - \overline{X\beta}^{j,k})$.

Residual defined as $var(\epsilon^{i,j,k})$+$2cov(\theta^{i} - \bar{\theta}^{j,k}, \epsilon^{i,j,k})$+ $2cov(X\beta^{i,j,k} - \overline{X\beta}^{j,k}, \epsilon^{i,j,k})$.

Table B.5: Industry Contributions to Between-Industry Variance Growth, Top 10 and Bottom 3 Industries (Females)

NAICS	Industry Title	Average $(\bar{y}^k - \bar{y})$	Change $(\bar{y}^k - \bar{y})$	Average (N_k/N)	Change (N_k/N)	Variance Contribut	Var Cont % of Total
7225	Restaurants & Other Eating Places	-.6374	-.0102	.0558	.0217	.0095	14.2%
6221	General Medical & Surg Hospitals	.3585	.1278	.0769	.0053	.0077	11.6%
5511	Management of Companies	.4594	.2240	.0204	-.0005	.0041	6.1%
4529	Other General Merchandise Stores	-.4399	-.0245	.0173	.0172	.0037	5.5%
6241	Individual and Family Services	-.3329	-.1810	.0134	.0100	.0028	4.2%
4481	Clothing Stores	-.4756	-.2669	.0113	-.0004	.0028	4.1%
5191	Other Information Services	.7277	.6026	.0016	.0021	.0027	4.1%
5417	Scientific Research Services	.7125	.2723	.0070	-.0003	.0026	3.8%
5112	Software Publishers	.9685	.1810	.0036	.0011	.0023	3.4%
4451	Grocery Stores	-.3356	-.1347	.0262	-.0010	.0022	3.3%
5179	Other Telecommunications	.5606	-.0312	.0020	-.0028	-.0010	-1.4%
3341	Computer Manufacturing	.8665	.2063	.0036	-.0030	-.0010	-1.5%
5171	Wired Telecommunications Carriers	.5966	.0245	.0091	-.0070	-.0022	-3.3%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.
(2) $(\bar{y}^k - \bar{y})$ is the industry's quarterly LN(real annual earnings) average relative to the economy average.
(3) (N_k/N) is the industry's employment share.
(4) Industry k's contribution to between-industry variance growth is $\Delta\left(\frac{N_k}{N}\right)[(\bar{y}^k - \bar{y})^2]$, which sums over industries to .0670.

Table B.6: Industry Contributions to Between-Industry Variance Growth, Top 10 and Bottom 3 Industries (Females)

NAICS	Industry Title	Variance Contribut	Segregat	Pay Premium	Sorting	Shift Share Earnings	Shift Share Emp Share
7225	Restaurants & Other Eating Places	14.2%	43.5%	9.6%	46.9%	7.6%	92.4%
6221	General Medical & Surg Hospitals	11.6%	9.7%	31.8%	58.5%	90.9%	9.1%
5511	Management of Companies	6.1%	49.7%	7.9%	42.4%	102.6%	-2.6%
4529	Other General Merchandise Stores	5.5%	34.5%	17.8%	47.7%	10.1%	89.9%
6241	Individual and Family Services	4.2%	43.6%	11.1%	45.3%	57.6%	42.4%
4481	Clothing Stores	4.1%	22.7%	26.5%	50.8%	103.8%	-3.8%
5191	Other Information Services	4.1%	26.0%	23.7%	50.2%	52.9%	47.1%
5417	Scientific Research Services	3.8%	52.8%	3.6%	43.5%	106.1%	-6.1%
5112	Software Publishers	3.4%	39.3%	13.6%	47.1%	55.4%	44.6%
4451	Grocery Stores	3.3%	25.8%	22.5%	51.7%	105.4%	-5.4%
5179	Other Telecommunications	-1.4%	13.5%	39.7%	46.8%	7.4%	92.6%
3341	Computer Manufacturing	-1.5%	7.0%	40.9%	52.1%	-130.4%	230.4%
5171	Wired Telecommunications Carriers	-3.3%	9.9%	45.3%	44.8%	-11.9%	111.9%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

(2) Industry k's contribution to between-industry variance growth is $\Delta[\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2]$, which sums over industries to .0670.

(3) The shift share term for changing employment share is $\left[\overline{\left(\bar{y}^k - \bar{y}\right)^2} * \Delta\left(\frac{N_k}{N}\right) / \Delta[\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2]\right]$. The shift share term for changing earnings is $\left[\overline{\left(\frac{N_k}{N}\right)} * \Delta\left(\bar{y}^k - \bar{y}\right)^2 / \Delta[\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2]\right]$.

Table B.7: Industry Contributions to Between-Industry Variance Growth, By Variance Contribution (Females)

Contribution to Variance		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
>5%	4 industries	.0251	37.4%	32.7%	17.4%	49.9%	49.2%	50.8%
1% - 5%	24 industries	.0398	59.4%	36.4%	16.2%	47.4%	80.6%	19.4%
0% - 1%	75 industries	.0143	21.3%	45.7%	11.9%	42.5%	107.1%	-7.1%
=0%	151 industries	.0002	0.3%					
<0%	47 industries	-.0124	-18.5%	19.6%	33.5%	46.8%	-22.8%	122.8%
		.0670	100.0%	39.8%	12.1%	48.1%	96.8%	3.2%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

Table B.8: Industry Contributions to Between-Industry Variance Growth, By Variance Contribution and Average Earnings (Females)

		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
28 industries with variance contribution >1%								
High-paying	16 industries	.0371	55.4%	33.6%	17.3%	49.1%	89.8%	10.2%
Low-Paying	12 industries	.0278	41.5%	36.9%	15.8%	47.2%	39.9%	60.1%
273 industries with variance contribution ≤1%								
High-paying	159 industries	.0016	2.4%					
Low-Paying	114 industries	.0005	0.7%					
		.0670	100.0%	39.8%	12.1%	48.1%	96.8%	3.2%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

Table B.9: Industry Contributions to Between-Industry Variance Growth, By Variance Contribution and Average Earnings (Females)

		Average $(\bar{y}^k - \bar{y})$	Change $(\bar{y}^k - \bar{y})$	Average (N_k/N)	Change (N_k/N)	Variance Contribut	Var Cont % of Total
28 industries with variance contribution >1%							
High Paying	16 industries	.4099	.1500	.2474	.0213	.0371	55.4%
Low-Paying	12 industries	-.4609	-.0820	.1851	.0693	.0278	41.5%
273 industries with variance contribution ≤1%							
High Paying	159 industries	.2881	.0580	.2774	-.0682	.0016	2.4%
Low-Paying	114 industries	-.3256	-.0093	.2902	-.0230	.0005	0.7%
						.0670	100.0%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

Figure B.7: Industry Contributions to Between-Industry Variance Growth, Males and Females