Industries and Increasing Inequality

By

John Haltiwanger, University of Maryland

Henry Hyatt, U.S. Census Bureau

James R. Spletzer, U.S. Census Bureau

December 21, 2020

Abstract

Two-thirds of the recent rise in U.S. males earnings inequality has occurred at the industry level. About ten percent of the over 300 4-digit NAICS industries account for nearly all of this increase in between-industry inequality. Using an AKM decomposition of earnings, the increase in between-industry inequality can be attributed primarily to changes in worker composition through sorting and segregation. For industries at the top of the earnings distribution, their contribution is dominated by rising inter-industry earnings differentials primarily accounted for by increased segregation of high paid workers. Among low-paying industries, expanding employment contributed to increasing inequality. The primary contributor for low-paying industries is increased sorting of low paid workers to these low-paying industries.

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). We thank Keith Bailey and Andrew Foote for helpful comments. 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 for our sample period from 1996 to 2018. Our contribution is to explore and emphasize the dominant role of industry effects in accounting for the rising between firm dispersion.[2] We also emphasize important distinctions between the evolution of earnings inequality for males and females.

We provide evidence that most (about 78 percent for males and 68 percent for females) of the rising between firm inequality is from rising inequality across more than 300 detailed 4-digit North American Industry Classification System (NAICS) industries. Moreover, a small fraction (slightly less than ten percent for both males and females) of industries account for virtually all the rising between industry dispersion. For males (females), just ten industries account for 70 (60) percent of the increase in between-industry dispersion while only accounting for 18 (23) percent of employment. The role of these industries in rising inequality varies depending on whether an industry tends to be high-paying or low-paying.

In a small number of high-paying industries, average worker pay soared. Increasing compensation in high-paying industries such as Software Publishers, Management of

[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.

Companies, Scientific Research Services, and Other Information Services was particularly important for rising inequality for both males and females. Increasing pay in the Computer Systems Design and Other Financial Services industries had a stronger contribution to earnings inequality among males than females. Analogous increases in the General Medical and Surgical Hospitals industry played a particularly important role for increasing inequality among females but not males.

At the same time as worker pay increased in high-paying industries, a small number of low-paying industries had flat or declining real compensation, but dramatic expansions of employment. Low-paying industries such as Restaurants and Other Eating Places and Other General Merchandise Stores have strong increases employment for both males and females. Males had a strong increase in employment in the Employment Services industry. Females had a strong increase in employment in the Clothing Stores and Individual and Family Services industries. The increasing employment share of these low-paying industries had an outsized role in terms of increasing in 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 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. We find some differences in the role of sorting vs. segregation based on whether an industry tends to be low-paying vs. high-paying. For low-paying industries, sorting plays an especially important role, as the lowest-paying industries are especially likely to employ workers with very low firm effects. Rising dispersion through sorting is less present among the highest-paying firms, where

we find more of a role for worker segregation. Rising dispersion in between industry average firm premia also play an important role: about 14 percent of the rising between industry dispersion is accounted for by rising dispersion in industry-level average firm premia.[3]

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.[4]

Song et. al. (2019) 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

[3] Our finding of an important role for rising firm premia contrasts with the findings of Song et. al. (2019). As we show below, this is partly due to differences in sample period. Song et. al. (2019) focus on the decomposition of rising earnings inequality from 1978 to 2013 while our sample period is 1996 to 2018. Song et. al. (2019) also present results for the sample period 1994 to 2013. For the 1994 to 2013 sample period, their AKM decomposition results yield a 15% contribution of rising dispersion in firm premia to rising overall earnings inequality, compared to a -1% contribution in the 1978 to 2013 sample period.

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

from Bureau of Labor Statistics (BLS) and Census processing. Both statistical agencies have strong incentives to have high quality detailed industry codes since they 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, it is structural change that yields changes in earnings and employment at the industry level. The firms in this small number of dominant of industries have common patterns of the changing organization of workers. In addition, we find that these industries have common patterns of changing firm premia.

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 between 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 supporting and connected role of rising between firm, within industry dispersion 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.[5] 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.[6] 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).[7]

We create our analytical dataset following Song et.al's sample restrictions. 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 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}).

[5] 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.

[6] 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.

[7] 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.

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).

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 Codes

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 $N \geq 1$ establishments with $N \geq M \geq 1$ 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. We see in the left panel of Figure 1 that the variance of male earnings was increasing from 1996 to 2012 but has been essentially flat during the 2012 to 2018 period.

The variance of female earnings was increasing from 1996 to 2015, with essentially no growth in the 2015 to 2018 period.

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:

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

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 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 male earnings increases from .7996 in 1996 to .9530 in 2018, an increase of roughly 15 log points. 29.7 percent of this increase is within firms (the red line in Figure 1) and 70.3 percent is between firms (the solid blue line in Figure 1). Looking at industry, 15.4 percent of the total increase in variance is within industries and 54.9 percent is between industries. This finding that the between industry component accounts for more than half of increasing male earnings inequality, and 78 percent of between firm inequality growth, is consistent with our earlier work in Haltiwanger and Spletzer (2020a, 2020b).

For females, 40.5 percent of increasing variance is within firms and 59.5 percent is between firms, with 20.6 percent within industries and 38.9 percent between industries. The within firm component of variance, which reflects the variance of individual earnings around the

firm mean, accounts for more growth for females than for males. Firm and industry contribute less to variance growth for females than for males.

To think ahead to the Song et.al (2019) variance decomposition, Tables 1 and 2 present the same variance decomposition using 7-year intervals rather than single years – Table 1 shows results for males and Table 2 shows results for females. Looking at males, the variance of earnings increases from .8358 in the first interval (1996-2002) to .9621 in the third interval (2012-2018). This .1263 increase is decomposed into .0196 within firms (15.5 percent), .0239 within industries (18.9 percent), and .0828 between industries (65.6 percent). Looking at females, the variance increases from .6676 in the first interval to .8068 in the third interval, with 28.6 percent within firms, 23.2 percent within industries, and 48.2 percent between industries.

We note that 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: looking at males in Table 1, 63.4 percent of variance in the first interval is within firms, 60.7 percent in the second interval, and 57.1 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 is relatively stable over time (for males, 15.3 percent in the first interval and 15.8 percent in the third interval). It is the between industry component of variance that is growing over time, from 21.3 percent in the first interval to 27.1 percent in the third interval.

II.D Earnings Percentiles

Figure 1 and Table 1 demonstrate that 4-digit NAICS industry accounts for over half of the growth of male 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.[8] In our analytical sample, comparing the first and the third intervals, male annual earnings declined by more than five log points for the first 38 percentiles, and declined for the first 61 percentiles. However, earnings at the top increased substantially. Earnings in the top 24 percentiles increased by more than five log points, and earnings in the top 14 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

[8] 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.

$(\bar{y}_t^k - \bar{y}_t)$. 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 $\Delta(y_t^{i,j,k} - \bar{y}_t^{j,k})$, the red line is the firm component $\Delta(\bar{y}_t^{j,k} - \bar{y}_t^k)$, and the blue line is the industry component $\Delta(\bar{y}_t^k - \bar{y}_t)$. 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. This general pattern of industry is the same for both males and females. 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 firm in Figure 2. For males, the firm component $\Delta(\bar{y}_t^{j,k} - \bar{y}_t^k)$ has essentially no contribution to the changing earnings distribution for the first 85 percentiles. The absolute value of the red line is less than .015 for each of the first 85 percentiles. From the 86th to the 99th percentiles, the males' firm component increases monotonically to a value of .080 for the final percentile.

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-

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

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

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).

IV. Person, Firm, and Covariance Effects

Tables 3 and 4 exploit the AKM decomposition of earnings to decompose rising earnings inequality using the person, firm, and covariance effects for males and females, respectively. The third column of each table presents the results using the LEHD data for the contribution of each of the relevant terms to the change in inequality from our first to last subperiods (1996-02 to 2012-18). For purposes of comparison, the first column presents results from the Song et. al. (2018) decomposition for the changing dispersion in earnings from the sub-period 1980-86 to 2007-13 (these are the results they emphasize). The second column presents their results for the decomposition using the change in dispersion of earnings for the subperiods that most closely overlap with ours (1994-2000 to 2007-13).

The results from the second and third columns are very similar for both males and females. Rising between firm dispersion dominates the rise in overall earnings inequality. From the mid-to-late 1990s to the most recent period, Song et.al find that 86.5 percent of male variance

growth is between firms, which is very similar to our result of 84.5 percent. For females, Song et.al find that 73.4 percent of variance growth is between firms and we find 71.4 percent.

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 5 (males) and Table A.1 (females). Looking first at variance growth of males from the mid-to-late 1990s to the most recent period, sorting contributes 37.5% in the Song et.al analysis and 35.3% in our LEHD data. Segregation contributes 35.5% in the Song et.al analysis and 37.3% 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 first column (Song et.al's variance growth for males from the early 1980s to the most recent period), with the notable exception of rising dispersion of firm premia.

The estimates for females are in Table A.1, and largely parallel the estimates for males. In the time period from the mid-1990s to the mid-2010s, sorting and segregation are the two largest components of between firm variance growth for females (29 to 33 percent), with the firm pay premia positive but smaller (8 to 12 percent).

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.2 in the appendix presents all the specific components for males while Table A.3 presents the specific components for females. 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 Tables 6 and 7 for males and females, respectively.

For males, 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 6 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 7 presents analogous results for females. We find that 47.9% (23.1% of 48.2%) of between industry dispersion is accounted for by increased sorting of high wage workers into industries with high average firm premia, 40.0% is accounted for increased segregation of workers by person effects across industries, and 12.0% is accounted for by increased dispersion in industry-level average firm premia across industries.

VI. Understanding the Dominant Role of Industry

VI.A Males

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 $\Delta var(\bar{y}^k - \bar{y})$, which is empirically estimated as:

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

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 8 and 9 provide statistics and decompositions for the top ten industries contributing to the rising between industry dispersion for males, along with the top three industries with the largest (in absolute value) drag on increasing earnings dispersion. Columns 1 and 2 in Table 8 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 variance growth (.0829 for males),

The top ten contributing industries for males account for 70% of the increase in between industry dispersion while only accounting for 18% of total employment.[9] As seen in Table 8, six of these ten industries are high-earnings industries with increasing earnings, and four are low-earnings industries with flat or declining earnings. The high-earnings industries in the top ten for males are Other Information Services, Software Publishers, Computer Systems Design,

[9] The calculations of employment shares for the top industries is 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.

Sorting and segregation play important roles in all the top ten contributing industries (Table 9). 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 premia range from 1.3% in Scientific Research Services to 24.1% in Other Information Services.

Whereas Tables 8 and 9 shows the top 10 and bottom 3 industries ranked by their contribution to increasing earnings dispersion, Table 10 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 (less than 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 50 percent of the 4-digit industries each of which contribute

zero percent,[10] about 20 percent that each contribute between zero percent and one percent, and about 20 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 10 with a negative contribution are only a modest drag on rising earnings inequality. Still the industries with the most negative contribution (see Table 8) 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.

The shift share estimates in Table 9 show that for the top ten industries, both changing inter-industry earnings differentials and changing employment contribute to rising between industry dispersion.[11] The magnitude of the shift share contributions vary substantially by industry. For many of the low-earnings industries in this top 10 list, most of contribution is from shifting employment towards these low-earnings industries (restaurants, employment services, and other general merchandise stores each have employment shift share estimates exceeding 70 percent). For many of the high-earnings industries in this top 10 list, rising inter-industry earnings differentials account for a larger share of the contribution.

We extend this shift share analysis by looking at the contribution of changing inter-industry earnings differentials and changing employment by the industry's location in the

[10] These 148 industries classified as zero percent each contribute less than .00004145 in absolute value to the growth in between industry variance growth (.00004145/.0829 = .00005, which is the threshold between 0.0 percent and 0.1 percent when rounding to one significant digit after the decimal point).

[11] 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 $\left[\overline{\left(\frac{N_k}{N}\right)} * \Delta\left(\bar{y}^k - \bar{y}\right)^2 / \Delta\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2\right]$.

earnings distribution. These estimates in Table 11 show, not surprisingly, that much of the increasing earnings dispersion occurs in the tails of the earnings distribution. 33 industries have average earnings that are 50 log points higher than the average, and these industries account for 35.3 percent of increasing between industry variance. At the other end of the distribution, 39 industries have average earnings that are 50 log points lower than the average, and these industries account for 40.8 percent of increasing between industry variance. At the top of the industry earnings distribution, earnings changes rather than employment changes account for all (105.6 percent) of the increase. At the bottom of the earnings distribution, employment changes account for most of the increase (75.2 percent).

An interesting feature of the results in Table 11 is that for the top earnings industries the contribution to rising between industry dispersion is accounted for almost completely by increased sorting and segregation between industries. For the bottom earnings industries, increased sorting and segregation are still dominant but rising dispersion in between industry firm-level pay premia is important as well.

These statistics in Table 11 help with the intuition of understanding increasing between industry dispersion. Industries at the top of the earnings distribution are characterized by increasing earnings, with this dominating any increases in employment. Industries at the bottom of the earnings distribution are characterized by increasing employment, with this effect substantially larger than any declines in real earnings.

In comparing the patterns in Table 11 with Tables 8 through 10, it might at first glance seem that Table 11 does not support the inference that increases in between industry dispersion is accounted for by a relatively small share of industries. This interpretation is incorrect. Taken collectively Tables 8 through 11 highlight that the industries that dominate rising between

industry dispersion are in the tails of the earnings distribution. However, being in the tails of the earnings distribution is not sufficient for an industry to make a large contribution to rising between industry dispersion. We illustrate this point in Table 12 where we report relevant statistics for the top 10 industries in terms of positive and negative inter-industry earnings differentials on average. Only three of the top ten contributing industries in Table 9 are in the list of top earnings industries, and each of these three industries is characterized by large increases in employment. An additional three of the top earnings industries are in the top ten percent of contributing industries. These high earnings industries have increasing earnings differentials but are not in the top ten contributing industries because their employment shares are either small and/or are not increasing over time.

Also, the top ten earnings industries includes the Computer Manufacturing industry that has a significant decline in employment, which contributes negatively to between industry dispersion. This pattern for Computer Manufacturing holds more generally in high tech manufacturing. Using the Hecker (2005) definition of high tech industries (STEM intensive industries), high tech manufacturing industries are all high earnings industries that collectively contribute a modest negative drag on between industry dispersion. The high tech industries outside of manufacturing are, in contrast, high earnings industries that contribute substantially to rising between industry dispersion. The primary distinction between the high tech manufacturing and high tech non-manufacturing is that the share of employment is declining in the former and rising in the latter.

Similar remarks apply to the bottom ten industries in terms of earnings differentials. Only two of the bottom ten earnings industries in Table 12 are among the top ten contributing industries in Table 8. These two industries, Restaurants and Other Eating Places and

Employment Services, both have large employment shares and large increases in employment. While not all low earnings industries are contributing substantially to rising between industry dispersion, the low earnings industries that are in the top ten percent of contributing industries are concentrated in Retail Trade and selected service industries.

An interesting pattern in Table 11 that the relative contribution of sorting and segregation varies between high and low paying industries. The high paying industries have a greater relative contribution of segregation compared to sorting. In contrast the low paying industries have the opposite. Given Table 12, it is useful to characterize the high and low paying industries by whether they exhibit increasing or decreasing earnings and employment. Table 13 explores these relationships for the top ten percent of industries that account for virtually all of the increase in between industry dispersion in earnings. Interestingly, only four of the possible eight cells are populated in the top ten percent of industries. Moreover, almost all of the increase in between industry dispersion is accounted for by three of the cells. Specifically, the cells with high average earnings and earnings increases, and the cell with low average earnings, declining earnings, and increasing employment. For these industries that account for virtually all of the increase in between industry dispersion, we find the pattern suggested by Table 11. That is, amongst the industries that account for most of the increase in between industry dispersion, increasing segregation is relatively more important than increased sorting for the high earnings industries. In contrast, for the low earnings industries that dominate the rise in between industry dispersion, increased sorting is relatively more important.[12]

To sum up, the top ten percent of contributing industries account for virtually all of the increase in between industry dispersion for males. The top ten percent tend to be in the tails of

[12] In unreported results, we find similar patterns for the top ten industries and for all 301 industries.

the earnings distribution. The top contributing industries with high earnings exhibit increasing earnings differentials. The top contributing industries with low earnings exhibit increasing employment. Increased sorting and segregation accounts for most 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 high-paying industries, increasing segregation is relatively more important than increased sorting. The opposite holds for the dominant low-paying industries.

VI.B Females

Analogous results are reported in Tables 14-17 for females. Many of the patterns are similar quantitatively and qualitatively for females. This is seen in Figure A.1, which plots the between industry variance contributions for all 301 industries for males and females. The unweighted correlation of the between industry variance contributions for males and females is 0.76.

The top ten percent of industries contribute almost 100 percent of the total between industry variance for females (Table 14) while accounting for only 43 percent of total (female) employment. Seven of the top ten industries for females overlap with the top ten industries for males. The industries that do not overlap are, for females, the low-earnings industries Clothing Stores and Individual and Family Services, both of which exhibit substantial declines in inter-industry earnings differentials. Individual and Family Services has a substantial increase in its employment share while Clothing Stores exhibits a modest decrease. The remaining non-overlapping industry for females is General Medical and Surgical Hospitals which is a high-earnings industry that exhibits a substantial increase in inter-industry earnings differentials and a

modest increase in employment share. As with males, the results for females show (Table 17) that rising inter-industry earnings differentials tend to be more important for the high-earnings industries while shifting employment tends to be more important for the low-earnings industries. In addition, increased segregation is relatively more important in high-paying industries while increased sorting is relatively more important in low-paying industries.[13]

VII. Between Firm, Within Industry

The increase in between firm, within industry variance of earnings plays an important supporting role accounting for about 20% of the overall increase in the variance of earnings for both males and females (tables 6 and 7). Interestingly, a relatively small share of industries also account for most of the between firm, within industry component.[14] Moreover, the industries that dominate the between industry component also tend to dominate the between firm, within industry component.

The top ten contributing industries to the between firm, within industry component for males are reported in Table 18.[15] These industries alone contribute 65% to the between firm, within industry component while accounting for only 17% of employment. The top ten percent of contributing industries account for more than 100 percent of the between firm, within industry component.[16] Four of the top ten industries in Table 18 are also among the top ten industries (for

[13] This is evident in Table 17 and Table A.4.

[14] The between firm within industry variance growth is $\Delta var\left(\bar{y}_t^{j,k} - \bar{y}_t^{k}\right)$, which is empirically estimated as $\sum_{k=1}^{301} \Delta\left(\frac{N_k}{N}\right)\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k} - \bar{y}^{k}\right)^2$. We define industry k's contribution to between firm within industry variance growth as $\Delta\left(\frac{N_k}{N}\right)\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k} - \bar{y}^{k}\right)^2$.

[15] Tables A.5 and A.6 provide more detail about the between firm, within industry component for males analogous to Tables 8 and 10 for the between industry component.

[16] See Table A.6. For the between firm, within industry component, the industries that negatively contribute account for -51%. Each of the negatively contributing industries have only a modest negative contribution as can be seen in Table 18 with the industry making the largest negative contribution, computer manufacturing, only accounting for -.0006, which is substantially smaller than the most positively contributing industries.

the between industry component) in Table 8. 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 employment-weighted correlation of the between industry and between firm, within industry contributions across the full set of 4-digit industries is 0.68 (the unweighted correlation is 0.59).

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 8). 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. For the between firm, within industry component, sorting, segregation and rising firm premia all play important roles.

For females, the top ten contributing industries account for about 60% of the between firm, within industry component (Table 19) while accounting for only 27% of employment. In addition, the top ten percent of contributing industries account for about 100% of the between firm, within industry component.[17] As with males, there is considerable overlap in the industries that contribute the most to the between industry and the between firm, within industry component. Comparing Table 19 and Table 14, there are two common 4-digit industries in the top ten of both but nine of the top ten share common 2-digit industries. Moreover, the

[17] See Table A.8.

employment-weighted correlation of the between industry and between firm, within industry contributions across the full set of 4-digit industries is 0.48 (the unweighted correlation is 0.53).

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. Most of the rising between firm inequality is accounted for by rising between industry dispersion in earnings. 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 less than ten percent of the 301 detailed 4-digit NAICS industries account for about 100% of the rising between industry dispersion (for both males and females). Even just the top ten contributing industries account for most of the rising between industry variance of earnings (70 percent for males and 60 percent for females). The top ten contributing industries for both males and females overlap considerably. The top ten industries 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, the contribution is dominated by rising inter-industry earnings differentials. For industries at the bottom of the earnings distribution, the contribution is dominated by shifts in employment to these very low-earnings industries. For industries at the top of the earnings distribution, increased sorting and segregation between industries dominates with increased segregation being relatively more important. For industries at the bottom of the earnings distribution, increased dispersion in between industry firm premia also plays an important supporting role. Increased sorting is relatively more important for the rising dispersion from the industries at the bottom of the earnings distribution.

The dominance of industry effects also shows up in the supporting role of the between firm, within industry component. A relatively small share of industries account for the between firm, within industry component. Moreover, the correlation between industries with large contributions to the between industry component and large contributions to the between firm, within industry component is positive and large (especially on an employment-weighted basis). This implies that the industries that dominate the between industry restructuring also tend to dominate the between firm, within industry restructuring.

The top contributing industries to rising between industry dispersion are in the high tech service industries like Software Publishers, Other Information Services, Scientific Research Services, and Computer Systems Design, retail trade industries like Grocery Stores, General Merchandise and Clothing Stores, the Restaurants and Other Eating Places industry, the Employment Services industry, and in health care industries like General Medical and Surgical Hospitals. The absence of goods producing industries is notable. Indeed, we find evidence that high tech manufacturing industries (e.g., Computer Manufacturing) are high earnings but a drag on rising earnings inequality.

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. 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. 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 low-paying 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 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.[18]

[18] 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

The most direct comparison of these findings with ours is using between industry dispersion from the AKM firm-level premium component. The latter abstracts from person effects as well as the covariance between the firm-level and person-level effects. Our approach estimating AKM over different sub-periods (as in Song et. al. (2019)) permits considerable flexibility in abstracting from these person and covariance effects. As discussed above, while most of the rising between industry dispersion in earnings is from increased sorting and segregation, we also find rising between industry dispersion in firm pay premia. Further investigation of this stark difference in findings from the CPS and our findings using administrative matched employer-employee data is needed. We note, however, that it has long been recognized that there are systematic limitations of industry codes in household surveys (see Mellow and Snider (1983) and Dey et. al. (2010)).[19] 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.

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.

[19] 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., 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.

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.

Hecker, Daniel E. 2005. "High-technology employment: a NAICS-based update." *Monthly Labor Review*, pp. 57-72.

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.

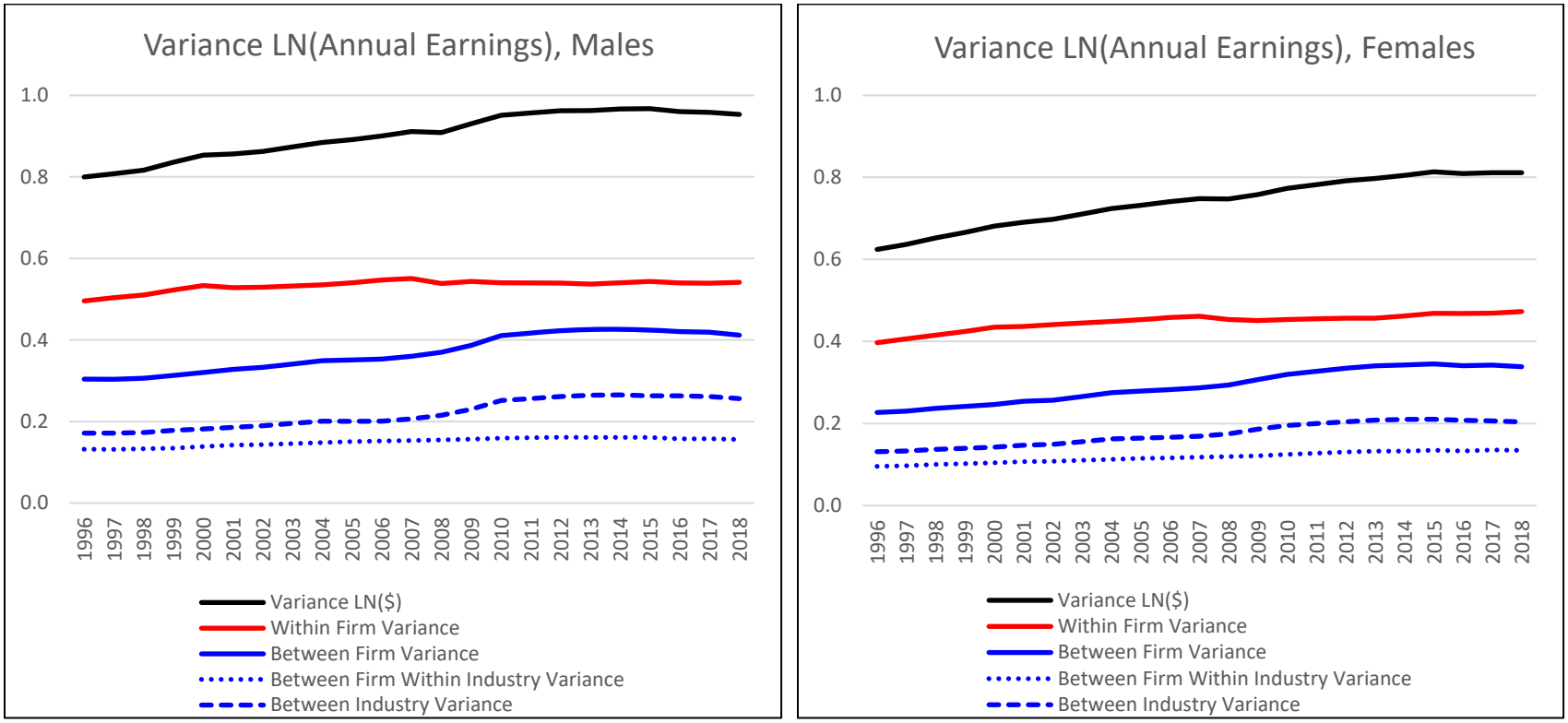
Mellow, Wesley, and Hal Sider. 1983. "Accuracy of response in labor market surveys: Evidence and implications." *Journal of Labor Economics* 1, no. 4:331–44.

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, 1996-2018



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

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

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

Table 1: Variance Decomposition, 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)	470.0	460.0	466.0	
Number of industries	301	301	301	

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

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

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

Table 2: Variance Decomposition, Females

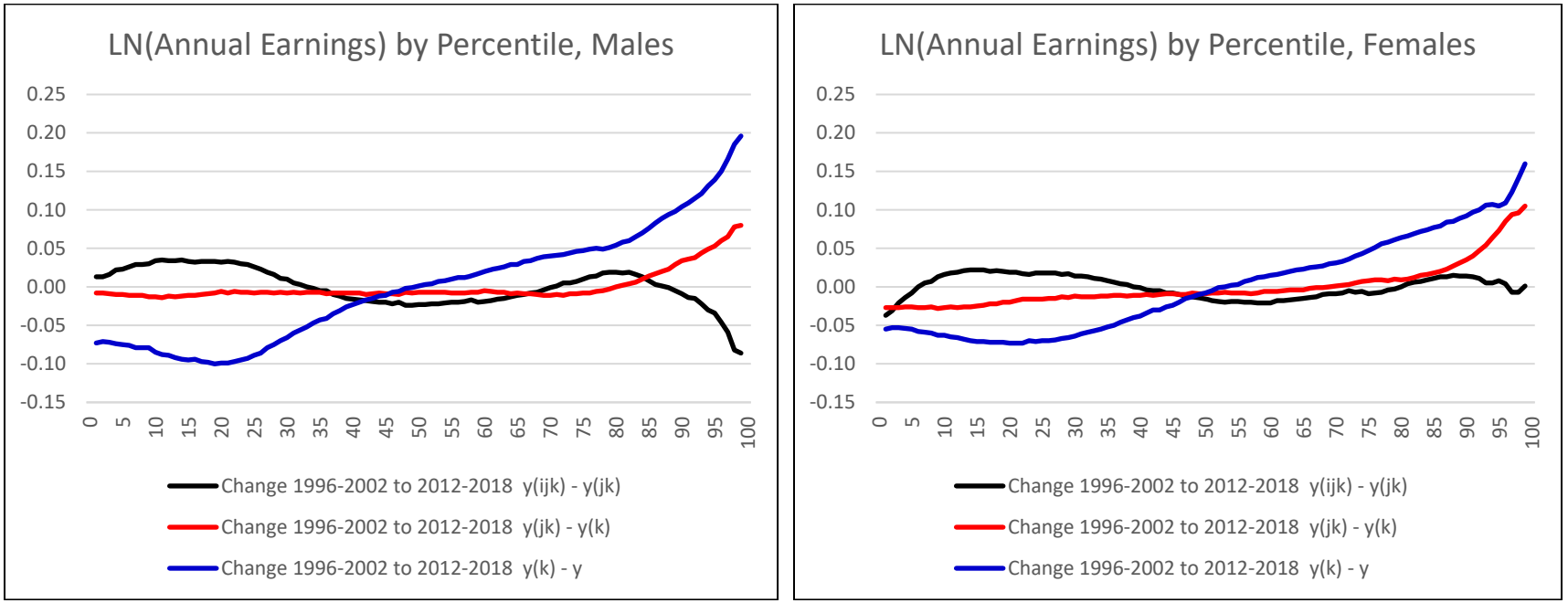
	1996-2002	2004-2010	2012-2018	Growth 1996-2002 to 2012-2018
Total $var\left(y^{i,j,k}\right)$	.6676	.7463	.8068	.1392
Person $var\left(y^{i,j,k} - \bar{y}^{j,k}\right)$	.4338	.4629	.4736	.0398
Firm $var\left(\bar{y}^{j,k} - \bar{y}^{k}\right)$	.0944	.1104	.1267	.0323
Industry $var\left(\bar{y}^{k} - \bar{y}_{t}\right)$	.1394	.1730	.2065	.0671
Person $var\left(y^{i,j,k} - \bar{y}^{j,k}\right)$	65.0%	62.0%	58.7%	28.6%
Firm $var\left(\bar{y}^{j,k} - \bar{y}^{k}\right)$	14.1%	14.8%	15.7%	23.2%
Industry $var\left(\bar{y}^{k} - \bar{y}_{t}\right)$	20.9%	23.2%	25.6%	48.2%
Sample Size (millions)	107.7	114.2	124.0	
Number of firms (thousands)	470.0	460.0	466.0	
Number of industries	301	301	301	

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

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

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

Figure 2: LN(Real Annual Earnings) by Percentile



Notes: (1) Left panel: males with annual real earnings >$3770 in EINs with 20 or more employees. Right panel: 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 also 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 3: Song et.al Variance Decomposition, Growth, Males

	Song et.al Growth 1980-1986 to 2007-2013	Song et.al Growth 1994-2000 to 2007-2013	LEHD Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.216	.096	.1263
Between firm $var(\bar{y}^{j})$	74.1%	86.5%	84.5%
$var(\bar{\theta}^{j})$	31.0%	30.2%	23.9%
$var(\Psi^{j})$	-1.4%	14.6%	11.8%
$var(\overline{X\beta}^{j})$	0.5%	-1.0%	1.8%
$2cov(\bar{\theta}^{j}, \Psi^{j})$	34.7%	33.3%	26.6%
$2cov(\bar{\theta}^{j}, \overline{X\beta}^{j})$	7.4%	6.3%	11.6%
$2cov(\Psi^{j}, \overline{X\beta}^{j})$	2.3%	4.2%	8.7%
Within firm $var(y^{i,j} - \bar{y}^{j})$	25.5%	13.5%	15.5%
$var(\theta^{i} - \bar{\theta}^{j})$	36.6%	26.0%	13.6%
$var(X\beta^{i,j} - \overline{X\beta}^{j})$	1.4%	-19.8%	3.0%
$var(\epsilon^{i,j})$	-8.3%	-10.4%	-1.0%
$2cov(\theta^{i} - \bar{\theta}^{j}, X\beta^{i,j} - \overline{X\beta}^{j})$	-3.2%	17.7%	2.0%
$2cov(\theta^{i} - \bar{\theta}^{j}, \epsilon^{i,j})$	-0.5%	-1.0%	-2.0%
$2cov(X\beta^{i,j} - \overline{X\beta}^{j}, \epsilon^{i,j})$	0.0%	0.0%	-0.2%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Song et.al estimates are computed from columns 1-10 of Table 4 of their 2019 QJE article.

Table 4: Song et.al Variance Decomposition, Growth, Females

	Song et.al Growth 1980-1986 to 2007-2013	Song et.al Growth 1994-2000 to 2007-2013	LEHD Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.221	.094	.1392
Between firm $var(\bar{y}^{j})$	53.8%	73.4%	71.4%
$var(\bar{\theta}^{j})$	21.7%	22.3%	20.3%
$var(\Psi^{j})$	-0.5%	11.7%	8.3%
$var(\overline{X\beta}^{j})$	1.4%	0.0%	2.9%
$2cov(\bar{\theta}^{j}, \Psi^{j})$	18.1%	26.6%	23.1%
$2cov(\bar{\theta}^{j}, \overline{X\beta}^{j})$	6.3%	6.4%	8.0%
$2cov(\Psi^{j}, \overline{X\beta}^{j})$	5.9%	6.4%	8.9%
Within firm $var(y^{i,j} - \bar{y}^{j})$	46.6%	27.7%	28.6%
$var(\theta^{i} - \bar{\theta}^{j})$	42.5%	38.3%	23.9%
$var(X\beta^{i,j} - \overline{X\beta}^{j})$	5.4%	-10.6%	6.1%
$var(\epsilon^{i,j})$	-5.0%	-16.0%	-2.7%
$2cov(\theta^{i} - \bar{\theta}^{j}, X\beta^{i,j} - \overline{X\beta}^{j})$	4.1%	16.0%	1.2%
$2cov(\theta^{i} - \bar{\theta}^{j}, \epsilon^{i,j})$	0.0%	0.0%	0.0%
$2cov(X\beta^{i,j} - \overline{X\beta}^{j}, \epsilon^{i,j})$	0.0%	0.0%	0.1%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Song et.al estimates are computed from columns 1-10 of Table A.9 of their 2019 QJE article.

Table 5: Song et.al Variance Decomposition, Growth, Males

	Song et.al Growth 1980-1986 to 2007-2013	Song et.al Growth 1994-2000 to 2007-2013	LEHD Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.216	.096	.1263
Between firm $var(\bar{y}^{j})$	74.1%	86.5%	84.5%
Segregation	38.9%	35.5%	37.3%
Pay Premium	-1.4%	14.6%	11.8%
Sorting	37.0%	37.5%	35.3%
Within firm $var(y^{i,j} - \bar{y}^{j})$	25.5%	13.5%	15.5%
Person effect	34.8%	23.9%	18.6%
Residual	-8.8%	-11.4%	-3.2%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Song et.al estimates are computed from columns 1-10 of Table 4 of their 2019 QJE article.
(3) 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 $2cov(\bar{\theta}^{j},\Psi^{j})+2cov(\Psi^{j},\overline{X\beta}^{j})$.
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 $var(\epsilon^{i,j})+2cov(\theta^{i} - \bar{\theta}^{j},\epsilon^{i,j})+2cov(X\beta^{i,j} - \overline{X\beta}^{j},\epsilon^{i,j})$.

Table 6: 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) 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 7: 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) 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 8: 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 9: 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	.0156	38.3%	13.7%	48.0%	25.9%	74.1%
5613	Employment Services	.0069	40.8%	13.2%	46.0%	0.0%	100.0%
4529	Other General Merchandise Stores	.0060	36.2%	16.3%	47.5%	28.9%	71.1%
5191	Other Information Services	.0053	27.0%	24.1%	48.9%	50.2%	49.8%
5112	Software Publishers	.0051	43.5%	11.6%	44.9%	41.1%	58.9%
5415	Computer Systems Design	.0050	63.2%	3.5%	33.3%	10.5%	89.5%
4451	Grocery Stores	.0048	29.8%	20.6%	49.7%	96.4%	3.6%
5511	Management of Companies	.0035	51.0%	7.3%	41.7%	104.3%	-4.3%
5239	Other Financial Investment Activity	.0030	48.2%	9.5%	42.3%	51.2%	48.8%
5417	Scientific Research Services	.0024	62.3%	1.3%	36.5%	116.3%	-16.3%
3345	Navigational Instruments Manuf	-.0009	14.9%	29.6%	55.4%	-73.9%	173.9%
5616	Investigation and Security Services	-.0009	3.2%	22.9%	74.0%	171.7%	-71.7%
3341	Computer Manufacturing	-.0011	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 $\Delta\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2$, which sums over industries to .0829.

(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 $\left[\overline{\left(\frac{N_k}{N}\right)} * \Delta\left(\bar{y}^k - \bar{y}\right)^2 / \Delta\left(\frac{N_k}{N}\right)\left(\bar{y}^k - \bar{y}\right)^2\right]$.

(4) The segregation, pay premium, and sorting terms are the industry's contribution to the terms in footnote 2 of Table 6.

Table 10: 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%	17.5%	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.

Table 11: Industry Contributions to Between-Industry Variance Growth, By Earnings Distribution, Males

Average Industry Earnings Relative to Grand Mean		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
>.5	33 industries	.0292	35.3%	51.8%	6.8%	41.3%	105.6%	-5.6%
.35 to .5	22 industries	.0082	9.9%	53.1%	2.3%	44.7%	110.8%	-10.8%
-.35 to .35	183 industries	.0056	6.7%	23.0%	55.5%	21.4%	113.5%	-13.7%
-.5 to -.35	24 industries	.0061	7.4%	22.3%	25.7%	52.1%	102.4%	-2.4%
< -.5	39 industries	.0338	40.8%	38.3%	15.0%	46.6%	24.8%	75.2%
		.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.

Table 12: Industry Contributions to Between-Industry Variance Growth, Top 10 and Bottom 10 Earnings Differentials 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)	Var Cont % of Total	Shift Share Earnings	Shift Share Emp Share
5112	Software Publishers	.9662	.1737	.0063	.0032	6.2%	41.1%	58.9%
2111	Oil and Gas Extraction	.9614	.1952	.0040	.0000	1.8%	102.5%	-2.5%
5231	Securities & Commodity Intermed	.9424	.1913	.0055	-.0010	1.2%	190.4%	-90.4%
5239	Other Financial Investment Act	.9091	.3418	.0025	.0017	3.7%	51.2%	48.8%
3341	Computer Manufacturing	.8653	.1698	.0063	-.0040	-1.4%	-161.5%	261.5%
4862	Pipeline Transport of Natural Gas	.8324	.0874	.0007	-.0001	0.0%	1244.7%	-1144.7%
3241	Petroleum & Coal Products Mfg	.8138	.2778	.0033	-.0008	1.2%	154.6%	-54.6%
5191	Other Information Services	.8112	.7077	.0023	.0034	6.4%	50.2%	49.8%
5232	Securities and Commodity Exch	.8086	.4448	.0001	.0000	0.1%	107.8%	-7.8%
4861	Pipeline Transport of Crude Oil	.7735	.1548	.0002	.0000	0.1%	82.6%	17.4%
8134	Civic and Social Organizations	-.6945	-.1297	.0014	.0000	0.3%	91.0%	9.0%
4854	School & Employee Bus Transport	-.7091	-.0181	.0012	.0000	0.0%	100.0%	0.0%
4512	Book Stores and News Dealers	-.7773	-.0041	.0007	-.0005	-0.4%	-1.4%	101.4%
4533	Used Merchandise Stores	-.7789	-.1446	.0007	.0004	0.5%	36.3%	63.7%
7225	Restaurants and Other Eating Places	-.8038	-.0577	.0436	.0179	18.8%	25.9%	74.1%
7224	Drinking Places (Alcoholic Bev)	-.8083	-.0155	.0011	.0001	0.1%	32.6%	67.4%
5613	Employment Services	-.8137	.0000	.0374	.0105	8.4%	0.0%	100.0%
1113	Fruit and Tree Nut Farming	-.8297	.1637	.0025	-.0001	-0.9%	91.7%	8.3%
7212	RV Parks and Recreational Camps	-.9053	.0256	.0002	.0001	0.1%	-14.6%	114.6%
1151	Support Activities for Crop Prod	-.9191	.1423	.0052	.0013	-0.3%	591.9%	-491.9%

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

Table 13: Top Ten Percent of Industries Contributing to Between-Industry Variance Growth, By High and Low Paying Industries and Earnings and Employment Changes, Males

Earn	Δ earn	Δ emp		Total	Total	Segregat	Pay Prem	Sorting	Earn Change	Emp Change
H	+	+	11 industries	.0292	35.3%	42.0%	17.2%	40.8%	48.2%	51.8%
H	+	-	8 industries	.0134	16.1%	50.1%	7.7%	42.1%	129.2%	-29.2%
H	-	+	0 industries							
H	-	-	0 industries							
L	+	+	0 industries							
L	+	-	0 industries							
L	-	+	8 industries	.0383	46.2%	36.1%	16.1%	47.9%	36.0%	64.0%
L	-	-	1 industry	.0008	1.0%	32.0%	18.8%	49.2%	109.5%	-9.5%
			28 industries	.0817	98.5%	40.4%	15.1%	44.4%	56.4%	43.6%

Notes: (1) Males with annual real earnings >$3770 in EINs with 20 or more employees.
(2) H=above average industry earnings; L=below average industry earnings. Δ earn + (-) indicates an industry with an increase (decrease) in real earnings per worker from 1996-02 to 2012-18. Δ emp + (-) indicates an industry with an increase (decrease) in employment shares over this same period.

Table 14: 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 15: 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	.0095	43.5%	9.6%	46.9%	7.6%	92.4%
6221	General Medical & Surg Hospitals	.0077	9.7%	31.8%	58.5%	90.9%	9.1%
5511	Management of Companies	.0041	49.7%	7.9%	42.4%	102.6%	-2.6%
4529	Other General Merchandise Stores	.0037	34.5%	17.8%	47.7%	10.1%	89.9%
6241	Individual and Family Services	.0028	43.6%	11.1%	45.3%	57.6%	42.4%
4481	Clothing Stores	.0028	22.7%	26.5%	50.8%	103.8%	-3.8%
5191	Other Information Services	.0027	26.0%	23.7%	50.2%	52.9%	47.1%
5417	Scientific Research Services	.0026	52.8%	3.6%	43.5%	106.1%	-6.1%
5112	Software Publishers	.0023	39.3%	13.6%	47.1%	55.4%	44.6%
4451	Grocery Stores	.0022	25.8%	22.5%	51.7%	105.4%	-5.4%
5179	Other Telecommunications	-.0010	13.5%	39.7%	46.8%	7.4%	92.6%
3341	Computer Manufacturing	-.0010	7.0%	40.9%	52.1%	-130.4%	230.4%
5171	Wired Telecommunications Carriers	-.0022	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 16: 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 17: Industry Contributions to Between-Industry Variance Growth, By Earnings Distribution, Females

Average Industry Earnings Relative to Grand Mean		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
>.5	34 industries	.0135	20.1%	62.0%	-6.2%	44.2%	181.6%	-81.6%
.35 to .5	30 industries	.0161	24.0%	29.1%	19.0%	51.8%	107.1%	-7.1%
-.35 to .35	197 industries	.0148	22.1%	34.5%	18.0%	47.5%	94.6%	5.4%
-.5 to -.35	23 industries	.0103	15.4%	32.3%	18.6%	49.1%	83.5%	16.5%
< -.5	17 industries	.0123	18.4%	42.3%	10.5%	47.1%	4.2%	95.8%
		.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 18: 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 19: Industry Contributions to Within-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
6216	Home Health Care Services	.0030	39.2%	14.5%	46.3%	37.7%	62.3%
5613	Employment Services	.0026	55.5%	9.0%	35.5%	96.0%	4.0%
6211	Offices of Physicians	.0024	52.5%	12.5%	35.0%	64.3%	35.7%
6241	Individual and Family Services	.0019	30.7%	27.6%	41.7%	27.9%	72.1%
6214	Outpatient Care Centers	.0018	38.0%	17.9%	44.0%	54.0%	46.0%
5416	Management Scientific Services	.0018	52.2%	14.4%	33.3%	20.9%	79.1%
6221	General Medical & Surg Hospitals	.0016	28.8%	25.8%	45.4%	85.9%	14.1%
7225	Restaurants & Other Eating Places	.0016	64.1%	15.4%	20.5%	26.4%	73.6%
5415	Computer Systems Design	.0013	51.6%	20.6%	27.8%	25.3%	74.7%
4461	Health and Personal Care Stores	.0012	41.5%	14.6%	43.9%	79.4%	20.6%
5111	Newspaper & Periodical Publishers	-.0006	23.6%	32.7%	43.6%	-58.2%	158.2%
3341	Computer Manufacturing	-.0006	46.4%	17.9%	35.7%	8.0%	92.0%
3344	Semiconductor Manufacturing	-.0006	21.1%	29.8%	49.1%	-111.2%	211.2%

Notes: (1) Females 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 .0324.

(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 $\left[\overline{\left(\frac{N_k}{N}\right)} * \Delta\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k}-\bar{y}^{k}\right)^2/\Delta\left(\frac{N_k}{N}\right)\sum_{j\in k}\left(\frac{N_J}{N_k}\right)\left(\bar{y}^{j,k}-\bar{y}^{k}\right)^2\right]$.

Appendix

Table A.1: Song et.al Variance Decomposition, Growth, Females

	Song et.al Growth 1980-1986 to 2007-2013	Song et.al Growth 1994-2000 to 2007-2013	LEHD Growth 1996-2002 to 2012-2018
Total $var(y^{i,j})$	.221	.094	.1392
Between firm $var(\bar{y}^{j})$	53.8%	73.4%	71.4%
Segregation	29.4%	28.7%	31.2%
Pay Premium	-0.5%	11.7%	8.3%
Sorting	24.0%	33.0%	32.0%
Within firm $var(y^{i,j} - \bar{y}^{j})$	46.6%	27.7%	28.6%
Person effect	52.0%	43.7%	31.2%
Residual	-5.0%	-16.0%	-2.6%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.
(2) Song et.al estimates are computed from columns 1-10 of Table A.9 of their 2019 QJE article.
(3) 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 $2cov(\bar{\theta}^{j},\Psi^{j})+2cov(\Psi^{j},\overline{X\beta}^{j})$.
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 $var(\epsilon^{i,j})+2cov(\theta^{i}-\bar{\theta}^{j},\epsilon^{i,j})+2cov(X\beta^{i,j}-\overline{X\beta}^{j},\epsilon^{i,j})$.

Table A.2: 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.

Table A.3: Industry Enhanced Variance Decomposition, Females

	1996-2002	2004-2010	2012-2018	Growth 1996-2002 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%
$var(\bar{\theta}^{j,k} - \bar{\theta}^{k})$	5.3%	5.6%	6.0%	9.5%
$var(\Psi^{j,k} - \bar{\Psi}^{k})$	3.9%	4.1%	3.7%	2.5%
$var(\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$	0.6%	0.4%	0.6%	0.6%
$2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}), (\Psi^{j,k} - \bar{\Psi}^{k})]$	3.5%	3.6%	4.1%	7.2%
$2cov[(\bar{\theta}^{j,k} - \bar{\theta}^{k}), (\overline{X\beta}^{j,k} - \overline{X\beta}^{k})]$	0.3%	0.5%	0.5%	1.7%
$2cov[(\Psi^{j,k} - \bar{\Psi}^{k}),](\overline{X\beta}^{j,k} - \overline{X\beta}^{k})$	0.5%	0.6%	0.7%	1.7%
Between industry $var(\bar{y}^{k})$	20.9%	23.2%	25.6%	48.2%
$var(\bar{\theta}^{k})$	4.8%	4.9%	5.8%	10.8%
$var(\bar{\Psi}^{k})$	4.8%	5.4%	5.0%	5.8%
$var(\overline{X\beta}^{k})$	0.4%	0.5%	0.8%	2.3%
$2cov(\bar{\theta}^{k}, \bar{\Psi}^{k})$	7.6%	8.5%	9.0%	15.9%
$2cov(\bar{\theta}^{k}, \overline{X\beta}^{k})$	1.3%	1.5%	2.1%	6.3%
$2cov(\bar{\Psi}^{k}, \overline{X\beta}^{k})$	2.0%	2.4%	2.9%	7.2%
Within firm $var(y^{i,j,k} - \bar{y}^{j,k})$	65.0%	62.0%	58.7%	28.6%
$var(\theta^{i} - \bar{\theta}^{j,k})$	41.9%	40.7%	38.8%	23.9%
$var(X\beta^{i,j,k} - \overline{X\beta}^{j,k})$	7.4%	5.3%	7.1%	6.1%
$var(\epsilon^{i,j,k})$	18.2%	16.1%	14.6%	-2.7%
$2cov(\theta^{i} - \bar{\theta}^{j,k}, X\beta^{i,j,k} - \overline{X\beta}^{j,k})$	-2.7%	-0.3%	-2.0%	1.2%
$2cov(\theta^{i} - \bar{\theta}^{j,k}, \epsilon^{i,j,k})$	0.1%	0.1%	0.1%	0.0%
$2cov(X\beta^{i,j,k} - \overline{X\beta}^{j,k}, \epsilon^{i,j,k})$	0.1%	0.1%	0.1%	0.1%

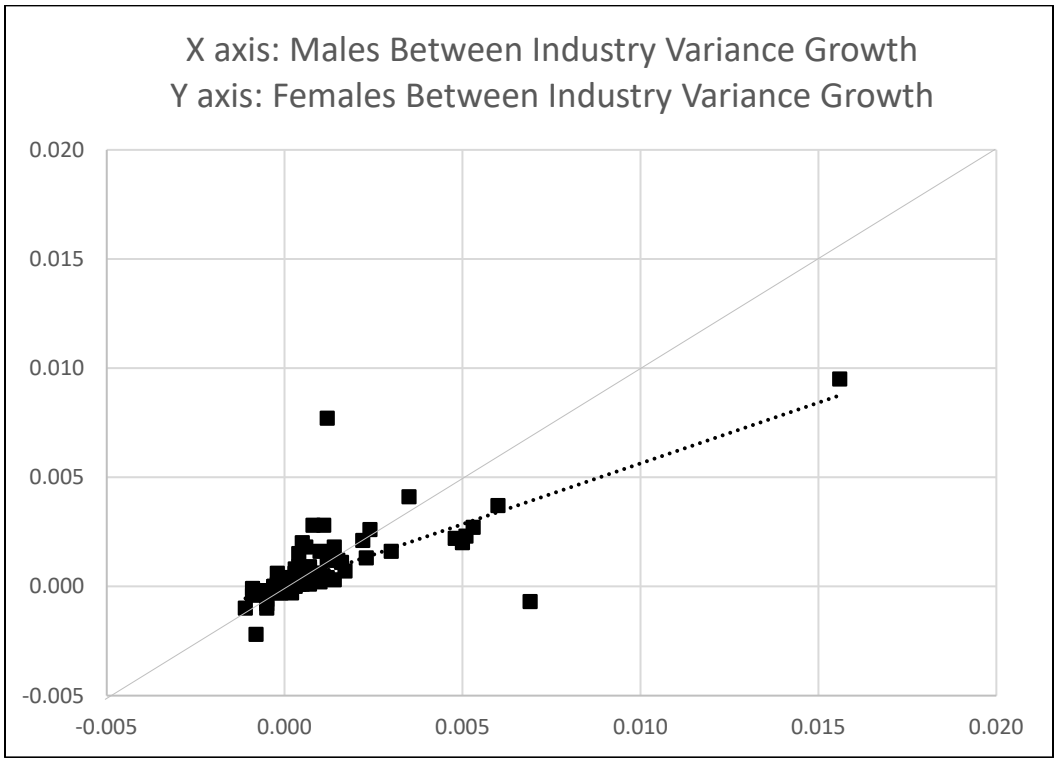
Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.

Table A.4: Top Ten Percent of Industries Contributing to Between-Industry Variance Growth, By High and Low Paying Industries and Earnings and Employment Changes, Females

Earn	Δ earn	Δ emp		Total	Total	Segregat	Pay Prem	Sorting	Earn Change	Emp Change
H	+	+	8 industries	.0207	30.9%	22.6%	24.7%	52.7%	73.6%	26.4%
H	+	-	7 industries	.0144	21.5%	45.7%	8.8%	45.4%	126.5%	-26.5%
H	-	+	1 industry	.0020	2.9%	59.2%	1.3%	39.5%	-8.9%	108.9%
H	-	-	0 industries							
L	+	+	0 industries							
L	+	-	0 industries							
L	-	+	10 industries	.0228	34.0%	39.7%	13.9%	46.4%	25.7%	74.3%
L	-	-	2 industries	.0050	7.5%	24.1%	24.7%	51.2%	104.5%	-4.5%
			28 industries	.0649	96.9%	48.3%	16.7%	35.0%	68.4%	31.6%

Notes: (1) Females with annual real earnings >$3770 in EINs with 20 or more employees.
(2) H=above average industry earnings; L=below average industry earnings. Δ earn + (-) indicates an industry with an increase (decrease) in real earnings per worker from 1996-02 to 2012-18. Δ emp + (-) indicates an industry with an increase (decrease) in employment shares over this same period.

Figure A.1: Between Industry Variance Contributions, Males and Females



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

Table A.5: 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.

Table A.6: 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%	37 industries	.0192	80.3%	45.1%	16.7%	38.2%	42.7%	57.3%
0% - 1%	80 industries	.0066	27.8%	53.2%	18.3%	28.5%	83.8%	16.2%
=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.

Table A.7: Industry Contributions to Within-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
6216	Home Health Care Services	-.3716	-.0465	.0149	.0069	.0030	9.1%
5613	Employment Services	-.5262	.0210	.0401	.0007	.0026	7.9%
6211	Offices of Physicians	.2368	.0999	.0281	.0077	.0024	7.4%
6241	Individual and Family Services	-.3329	-.1810	.0134	.0100	.0019	5.9%
6214	Outpatient Care Centers	.2850	.2200	.0104	.0078	.0018	5.7%
5416	Management Scientific Services	.3877	.0590	.0089	.0055	.0018	5.6%
6221	General Medical & Surg Hospitals	.3585	.1278	.0769	.0053	.0016	5.0%
7225	Restaurants & Other Eating Places	-.6374	-.0102	.0558	.0217	.0016	4.8%
5415	Computer Systems Design	.6324	-.0114	.0122	.0054	.0013	3.9%
4461	Health and Personal Care Stores	-.2063	-.0147	.0107	.0022	.0012	3.8%
5111	Newspaper & Periodical Publishers	.2570	.0878	.0064	-.0059	-.0006	-1.7%
3341	Computer Manufacturing	.8665	.2063	.0036	-.0030	-.0006	-1.7%
3344	Semiconductor Manufacturing	.3973	.2249	.0063	-.0050	-.0006	-1.8%

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

Table A.8: Industry Contributions to Within-Industry Variance Growth, By Variance Contribution, Females

Contribution to Variance		Total	Total	Segregat	Pay Premium	Sorting	Earn Change	Emp Change
>5%	7 industries	.0151	46.7%	43.3%	16.5%	40.2%	55.7%	44.3%
1% - 5%	25 industries	.0180	55.7%	50.3%	14.5%	35.2%	61.2%	38.8%
0% - 1%	82 industries	.0090	27.8%	46.1%	15.3%	38.6%	84.9%	15.1%
=0%	89 industries	-.0001	-0.2%					
<0%	98 industries	-.0097	-30.0%	32.0%	30.3%	37.7%	-30.0%	130.0%
		.0324	100.0%	51.7%	10.1%	38.2%	97.4%	2.6%

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