BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Employment Research and Program Development

Outsourcing, Occupationally Homogeneous Employers, and Wage Inequality in the United States

Elizabeth Weber Handwerker, U.S. Bureau of Labor Statistics

Working Paper 522
September 2022

Outsourcing, Occupationally Homogeneous Employers, and Wage Inequality in the United States

Elizabeth Weber Handwerker
U.S. Bureau of Labor Statistics

September 2022

Abstract: This paper develops measures of the occupational homogeneity of employers as indicators of outsourcing. Findings are threefold. First, wages are strongly related to occupational homogeneity, particularly for workers in low-wage occupations. Second, by some measures, workers—particularly those in lower-wage occupations—saw their employing establishments become more occupationally homogeneous during 2004-2019. Third, changes in the occupational homogeneity of workplaces are an important contributor to growing wage inequality among workers over the first half of this period. The growing sorting and segregation by occupation of workers into different employers is an important part of wage inequality growth.

I am grateful to seminar participants at the American Economic Association 2014 Annual Meetings, the University of Maryland, the Economic Policy Institute, the Federal Reserve Board of Governors, the Society of Labor Economists 2016 and 2017 Annual Meetings, the Center for Equitable Growth, the NBER 2018 Summer Institute, the BRIQ Institute, the Bureau of Labor Statistics, and the NBER Wage Dynamics in the 21st Century conference for comments, with special thanks to Matt Dey, Jesse Rothstein, Katharine Abraham, Andy Garin, David Autor, Erica Groshen, Sue Houseman, David Levine, Anne Polivka, Alex Bryson, David Piccone, David Weil, Nicole Fortin, Fernando Rios-Avila, and anonymous referees for particularly helpful comments and suggestions. This paper grew out of a project begun jointly with James Spletzer, and I thank him for his many contributions. Any opinions and conclusions herein are those of the author and do not necessarily reflect the views or policies of the U.S. Bureau of Labor Statistics. Research assistance from Lowell Mason and David Piccone is gratefully acknowledged.

I. Introduction

Growing inequality of wages, particularly between employers, has been a key feature of the labor market in recent decades. Many changes in the labor market have been examined as potential sources of this inequality growth—including the decline of manufacturing, the role of technology in replacing employer demand for routine work, and the increased potential for imported goods and services to replace domestic production. This paper examines an additional source of growing wage inequality: the changing distribution of occupations between employers as the organization of production changes, with employers retaining certain types of work within the workplace and outsourcing other work.

Much evidence shows that establishments play an important role in determining individual wages, beyond the role of individual workers' characteristics (Groshen 1991a, 1991b; Bronars and Famulari 1997; Abowd, Kramarz, and Margolis 1999; Lane, Salmon, and Spletzer 2007; Card, Heining, and Kline 2013). Several authors have used employer microdata to study growing variability in earnings in the U.S. from the mid-1970s to the early 2000s, and have found it due more to variation between establishments than to variation within establishments (Davis and Haltiwanger 1991; Dunne, Foster, Haltiwanger, and Troske 2004; Barth, Bryson, Davis, and Freeman 2016; Handwerker and Spletzer 2016; and Song, Price, Guvenen, Bloom, and von Wachter 2019),[1] while the increased sorting of high-paid workers to high-paying employers drives much of the growth in pay inequality between employers (Song, Price, Guvenen, Bloom, and von Wachter, 2019). The results in this paper show that occupational

[1] There is a large and growing literature on wage inequality growth in Europe, based on employee-employer linked data, including Card, Heining, and Kline (2013), who emphasize the role of increased worker sorting between employers in explaining wage inequality growth in Germany.

homogeneity—a specific form of worker sorting—is a key explanation for this growth in between employer wage inequality. More and more workers in high-wage occupations are employed in different workplaces from workers in other occupations, exacerbating differences in their pay.

The intersection of growing underlying wage inequality and the business environment in the United States can make it profitable for employers to focus on employing either low or high wage workers. Growing wage inequality among workers has arisen from such sources as the changing composition of the workforce and changing returns to education and experience,[2] the growing inequality within education and skill groups[3], and the differential impact of technology on the worker skill distribution[4]. As wages for different kinds of work became less equal, employers faced regulations requiring nondiscrimination across employees in the coverage of pension, health insurance and other benefits (EBRI 2009, Perun 2010),[5] increasing incentives to contract out work that pays very different wages from the work of other employees. Moreover, social norms may make it more acceptable for employers to contract out work rather than pay very different wages to employees doing different kinds of work (Weil 2014).

Other potential reasons for businesses to outsource work include increasing ability to smooth workload, economies of scale available to providers of specialized services (Abraham and Taylor, 1996), or a focus on "core competencies" enabled by technologies for specifying and

[2] Bound and Johnson 1992, Katz and Murphy 1992, Lemieux 2006
[3] Juhn, Murphy, and Pierce 1993, Katz and Autor 1999
[4] Juhn, Murphy, and Pierce 1993, Acemoglu 2002, Autor, Katz, and Kearney 2006, 2008
[5] Perun (2010) lists a variety of employment benefits which receive favorable tax treatment and are required to be available to low-wage as well as high-wage employees of each employer.

monitoring work done by outsiders (Weil 2014). However, to the extent that labor cost savings and avoidance of efficiency wages or rents for occupations with low wages in the labor market are key reasons for outsourcing, it can lead to employers specializing in high or low-wage work, and result in growing wage inequality between establishments. Goldschmidt and Schmeider (2017) show labor cost savings to be a primary reason for outsourcing in Germany, as outsourced workers lose firm-specific rents, while Drenik, Jäger, Plotkin, and Schoefer (2021) study this same phenomenon for the outsourcing of work to temp agencies in Argentina, and Bilal and Lhuillier (2021) model its impact in France. In three well-defined occupational categories, Goldschmidt and Schmeider find that losses of such firm-specific rents can account for 9% of all growth in German wage inequality from 1985 to 2008.

In U.S. data, direct measures of outsourcing are not generally available. Researchers have instead focused on particular industries or occupations associated with performing support tasks for other businesses. Dey, Houseman, and Polivka (2010) show a marked increase in various measures of outsourcing in recent years such as trends in temporary help or employment services. Estimates from several sources show these industries roughly doubling in size from 1992 to 2002. They also document an increase in the employment share of occupations associated with outsourced labor services, such as school bus and truck drivers in the transportation industry and accountants in the business services industry. Yet these measures only capture a fraction of outsourcing—that which occurs in these specific industries. Dube and Kaplan (2010) use individual-level data to show the impact of outsourcing on wages and benefits for janitors and guards, but again, their measures can only capture outsourcing of a narrow set of occupations.

This paper develops economy-wide measures of outsourcing in the United States, using the homogeneity of occupations by employer, as measured in the detailed microdata of the Occupational Employment and Wage Statistics Survey conducted by the Bureau of Labor Statistics. These measures distinguish between two types of outsourcing, which may have differing impacts on wage inequality. When businesses outsource work to avoid monitoring, hiring, or other costs for occupations in which they have less expertise, there will be less variety overall in the occupations they employ. However, when businesses outsource work to narrow the wage distribution of their employees, the variance of wages predicted from the particular set of occupations they employ will decrease. The impact of the changing distributions of occupations and of employer occupational homogeneity are compared with the effects of other changes in employer characteristics (industry, size, and location) on the overall distribution of wages.

There are three major findings. First, wages are related to the occupational homogeneity of establishments. Workers in more occupationally homogeneous establishments earn lower wages. This relationship holds even after controlling for workers' own occupations and observable characteristics of their employers and is strongest for workers in occupations typically paid lower wages. Second, from 2004 through 2019, the occupational homogeneity of employers increased by some measures, increasing unambiguously for workers in typically high-wage occupations, after controlling for other employer characteristics. Third, changes in the distribution of this occupational homogeneity are related to the growth in private-sector wage inequality observed in the data from 20014 through 2016. A substantial amount of the growth in ln(wage) variance, as measured in the OEWS data, can be attributed to the growing occupational homogeneity of establishments over this period. Both measures of employer homogeneity—one based on the distribution of occupations by wage levels, and the other a more functional measure of employer homogeneity that ignores wage differences among occupations—matter for growing wage inequality.

The paper is organized as follows: Section II describes measures of occupational homogeneity. Section III describes relationships between employer occupational homogeneity and employee wages. Section IV describes trends in measured occupational homogeneity of employers. Section V describes the impact of the changing distributions of occupation and the occupational homogeneity of employers on wage inequality over time. Section VI concludes.

II. Measuring the Occupational Homogeneity of Employers

I use the term "occupational homogeneity" to describe the variety of occupations employed at a place of business, separate from the tasks performed by individual employees (their occupations), the type of work done at the business (its industry) or the size of the business. Much scholarship on outsourcing (for example Dey, Houseman, and Polivka, 2010; and Erickcek, Houseman, and Kalleberg, 2003) examines particular occupations and particular industries. In contrast, occupational homogeneity is intended as a measure of the variation in work done in all businesses, through the full range of industries in the economy. This section defines two measures of occupational homogeneity and presents evidence showing that these measures are related to examples in the outsourcing literature.

The two measures of the occupational homogeneity of establishments are very different: (1) a measure involving the overall distribution of occupations, regardless of whether they are high or low paid, and (2) a measure that explicitly models the variation in wages of establishments due to the distribution of occupations employed.

The first measure of occupational homogeneity for establishment *j* at time *t* is constructed with a Herfindahl-Hirschman index of employment, *n*, in each occupation *k* within that establishment, normalized for the overall size of the establishment, N:

(1) $H_{jt} = \sum_{k=1}^{100} \left(\frac{n_{kjt}}{n_{jt}} \right)^2$ $\quad$ $Normalized\ H_{jt} = \frac{(H_{jt} - {}^1/_N)}{(1 - {}^1/_N)}$, or 0 if N = 1

This index uses the 100 minor occupational categories at the 3-digit level of the Standard Occupational Classification system.[6] It varies from 0 (equal representation of all occupations) to 1 (complete homogeneity). Increased occupational homogeneity at the establishment level by this measure indicates that employers are becoming more specialized, consistent with outsourcing work to other employers. Trends in this measure indicate whether establishments throughout the U.S. economy are becoming more homogeneous in the occupations they employ. However, this measure cannot distinguish between specializing in a few occupations typically paid very different wages, such as 29-1000 (Healthcare Diagnosing or Treating Practitioners)

[6] Handwerker and Spletzer (2016) studied this type of general occupational homogeneity with Herfindahl-Hirschman indices, using both the detailed 6-digit occupations of the Standard Occupational Classification System (829 categories) and the 2-digit major occupational categories of the Standard Occupational Classification System (22 categories), and found very similar time trends and relationships between occupational classification and wages with broad and detailed versions of this measure.

and 31-1100 (Home Health and Personal Care Aides; and Nursing Assistants, Orderlies, and Psychiatric Aides), or specializing in a similar number of occupations that are typically paid more similar wages.

In contrast, the second measure of occupational homogeneity is explicitly constructed to capture the similarity or dissimilarity of typical wages for the occupations employed at an establishment. It is the part of the variance of wages for each establishment that would be predicted from the establishment's distribution of employment by occupation, without using information on the actual wages paid at the establishment. Using average log wages for each minor occupational category in each time period, the log wage paid by employer *j* to worker *i* in occupation *k* at time *t* is estimated as $\widehat{w_{ijt}} = \overline{w_{kt}} + \varepsilon_{ijt}$, where $\overline{w_{kt}}$ is the mean log wage for all employees in occupation *k* at time *t* and ε_{ijt} is distributed normally, with mean 0 and standard deviation σ_k. From the occupational distribution of employer j at time t, the estimated mean log wage for j at t is estimated $\widehat{\widehat{w_{jt}}} = \frac{\sum_k \sum_{i \in k} \overline{w_{kt}}}{n_{jt}}$, where n_{jt} is the total employment for employer *j* at time *t*, and $i \in k$ denotes observations in which individual *i* has occupation *k*. Again, using only the distribution of occupations employed and the average wages of these particular occupations across all employers at time t, the predicted log wage variance for employer *j* at time *t* is

$$\widehat{V_{jt}} = \frac{\sum_i \left(\widehat{w_{ijt}} - \widehat{\widehat{w_{jt}}}\right)^2}{n_{jt}} = \frac{\sum_k n_{jkt}\left[\left(\overline{w_{kt}} - \widehat{\widehat{w_{jt}}}\right)^2\right]}{n_{jt}} + \frac{\sum_k n_{jkt}\sigma_{kt}{}^2}{n_{jt}}.$$

This predicted variance has two parts. The first is the variation in average log wages between the occupations employed in the establishment. The second is the average of within-occupation log wage variances for these occupations. Occupations with higher wages tend to have higher within-occupation wage variances. Thus, only the first part of the predicted log wage variance is the second measure of occupational homogeneity.

$$(2) \qquad \frac{\sum_k n_{jkt}\left[\left(\overline{w_{kt}} - \widehat{\overline{w_{jt}}}\right)^2\right]}{n_{jt}}$$

Both these measures are estimated with the microdata of the Occupational Employment and Wage Statistics (OEWS) Survey for the private sector in the United States for 2004 through 2019. These microdata record the number of employees by wage interval within detailed occupation categories for hundreds of thousands of establishments per year. The OEWS survey is designed to produce estimates of employment and wages in the United States for each detailed occupation, by geography and industry. It covers all establishments in the United States except for those in agriculture, private households, and unincorporated self-employed workers without employees. It is the only survey of its size and scope.

The OEWS collects data for a sample of about 200,000 establishments each November and each May. Sampled establishments are asked to report the number of employees in each occupation by wage interval. In using OEWS data to study wage inequality, it is important to understand that the OEWS data *cannot* measure inequality in the topmost percentiles of the wage distribution. Wages are reported to the OEWS in intervals. Earnings of individuals at the very top of the wage distribution are topcoded in the OEWS—the uppermost interval in the recent OEWS surveys is “$208,000 and over.” Averaged across all years, the uppermost interval contains roughly 1.3 percent of employment. Handwerker and Spletzer (2014) compare wage inequality levels and trends in these OEWS microdata with the wage inequality level and trends in the outgoing rotation group microdata of the CPS, which has been used in many of the most cited studies of wage inequality. They show the interval nature of wage collection in the OEWS has almost no impact on overall wage variance trends. Both this study and Dey, Handwerker, Piccone, and Voorheis (2022) show OEWS data broadly replicate CPS wage distribution levels and trends. Overall wage variances in each year are similar in the reweighted OEWS and CPS microdata until 2016. However, from 2016 to 2019, Dey, Handwerker, Piccone, and Voorheis show there was a more substantial wage variance decline in OEWS data than in CPS data.

The OEWS sample design uses 3 years, or 6 panels of data collection, to produce detailed published estimates of employment and wages. It is not designed to produce time series estimates of either employment or wages for any individual occupation, in part because of changes over time in occupational definitions. This paper uses OEWS microdata from November 2004 (collected from 2001 through 2004), November 2007 (collected from 2005 through 2007), November 2010 (collected from 2008 through 2010), November 2013 (collected from 2011 through 2013), November 2016 (collected from 2014-2016), and November 2019 (collected from 2017-2019). Various adjustments are made to occupations and industries to make them as consistent as possible throughout the period. Public sector employers are not included in this analysis because the establishment of employment is not part of OEWS data collection for portions of the public sector.

Establishments are the sampling units of the OEWS, and so this paper focuses on measures of occupational homogeneity at the establishment level. However, all the main results in this paper have been repeated with measures constructed at the Employer Tax-ID level (EIN), and results are shown in Appendix C.

As noted above, Dey, Handwerker, Piccone, and Voorheis show there was a substantial wage variance decline in OEWS data from 2016 to 2019 (following a plateau in wage variance from 2013 to 2016). Analyses in that paper show the wage convergence was due to increases in wages for lower-wage occupations, rather than to shifts in the occupational distribution. Because of these wage changes, the second measure of occupational homogeneity must be estimated using average wages for each occupation that are constant throughout the period. If this measure is estimated instead using period-specific average wages for each occupation, the increased wages for lower-wage occupations generate a large spurious reduction in this measure of occupational homogeneity between 2016 and 2019.

The Data Appendix contains more details about the OEWS. It also contains summary statistics for the data in this paper, including the composition of occupations and industries. The average worker has an inflation-adjusted wage of $16.55/hour (in $2000), or a ln(wage) of 2.59, and is observed in an establishment with a measured ln(wage) variance of 0.166. The average normalized Hirfindahl-Hirschman index for workers' establishments is 0.360, and the between-occupations portion of the average predicted variance of ln(wages) estimated from its workers' occupational composition is 0.103.

Table 1 compares the two measures of establishment-level occupational homogeneity for several occupation-industry groups studied as examples of outsourcing by Abraham and Taylor (1996); Dube and Kaplan (2010); Dey, Houseman, and Polivka (2010); Weil (2014); and Goldschmidt and Schmeider (2015): the entire food preparation and serving major occupational group, janitors, security guards, truck drivers, accountants, computer occupations, engineering occupations, and lawyers. Outsourcing of workers in these occupations means that they are employed in the specialty industries of food services, janitorial services, security guard services, truck transportation, accounting services, computer services, engineering services, or law offices, rather than the industry of the business to which they provide these services. Table 1 shows that for every single one of these example occupations or occupation groups, the normalized Herfindahl-Hirschman indices for employers of these workers (as defined in equation (1)) are higher, on average, indicating greater occupational homogeneity of employers, when they are employed in their specialty industry than when they are employed in other industries. Moreover, for every example occupation except lawyers (the smallest and highest paid of these examples), the partial predicted variances of wages based on the occupational distribution of their employers are lower, on average, indicating greater occupational homogeneity of employers, when they are employed in their specialty industry than when they are employed in other industries. Both measures of occupational homogeneity measures defined in this section—designed to measure outsourcing across all occupations and industries—indicate greater occupational homogeneity in the relevant industries to which workers are outsourced than in other industries, for all the individual occupations studied in the outsourcing case-study literature (except lawyers).

The final row of Table 1 examines occupations that are most dispersed across industries—the 11 minor occupational groups comprising the 20% of employment with the highest HHI values across detailed industries. These occupational groups include Top Executives (which seems quite unlikely to be outsourced), as well as ten others that seem possible to outsource to the Business Services Industry, such as Material Recording, Scheduling, Dispatching, and Distributing Workers and Information and Record Clerks. None of these eleven occupational groups are in the lowest-paid quintile of occupations. This group of occupations follows the same pattern as the example occupations from the outsourcing literature shown earlier in the Table, with higher concentration levels of their establishments (by either measure) in the Business Services Industry than in other industries, and lower wages in the Business Services Industry than in the other industries where they are employed.

Among the occupations in Table 1, all the low wage occupations (food preparation and service, janitors, and security guards) earn considerably lower average wages in outsourced specialty industries than in other industries. These example occupations are examples precisely because there are obvious industries to which they can be outsourced; most other occupations do not have such obvious industries for outsourcing. However, the advantage of the occupational homogeneity measures in this paper is that they can be measured for the employers of all occupations. The next section shows the relationship between occupational homogeneity and wages for all workers.

III: Relationships between Measured Occupational Homogeneity and Wages

This section uses regressions to describe the relationship between occupational homogeneity and wages, following the specification

(3) $$\ln(wage_{ijt}) = \alpha OccHomogeneity_{jt} + \beta X_{ijt} + \varepsilon_{ijt},$$

where $OccHomogeneity_{jt}$ measures occupational homogeneity for the employer j of individual *i* at in time *t*, and X_{ijt} are other observable characteristics of individual *i* (occupation) and employer *j* (industry, geography, and size) at time *t*. Results of this regression are shown in Table 2. The first row of this table gives estimates of the impact of occupational homogeneity on wages, α, with no additional variables (other than a fixed effect for each reference date). These estimates show that increased occupational homogeneity is associated with lower wages overall. The second row of Table 2 gives these estimates with all X_{ijt} variables added to the regression. These detailed controls reduce the magnitude of the relationship between occupational homogeneity and wages, α, but the estimates maintain the same sign and remain very significant.

Further rows of Table 2 repeat this analysis for subgroups of occupations. Occupations (at the 3-digit minor occupational category SOC level) are grouped by average wage into quintiles, with roughly equal total weighted employment in each quintile.[7] Appendix A lists the occupations of each quintile, while counts of the observations for each quintile are in the Data Appendix. The list of occupations in the lowest-paid quintile is a short one, because the occupations in this quintile, such as Food and Beverage Serving Workers, tend to be large. The list of occupations in the highest-paid quintile, such as Social Scientists, is much longer, because these occupations tend to be smaller.

The relationship between occupational homogeneity and wages, after controlling for own-occupation and employer characteristics, is generally stronger for workers in typically low- and middle-wage occupations than for workers in typically high-wage occupations. The

[7] To form quintiles, occupations are ranked by their average wages across all years. This grouping of occupations is quite stable over time.

relationship between the typical wage levels for a quintile of occupations and the wage coefficient of occupational homogeneity for the occupations in that quintile is not monotonic, with the largest wage coefficients for the quintile of occupations with the second-lowest typical wages.

There is one group of workers for whom greater occupational homogeneity—at least as measured by the predicted variance of wages between occupations—is associated with substantially *higher* wages, once own-occupation and employer characteristics are taken into account. These are the workers in the highest paid quintile of occupations. This is consistent with the model of Bilal and Lhuillier (2021), in which the outsourcing of lower-paid work is associated with greater demand—and higher wages—for work in higher-median-wage occupations. It is also consistent with the notions that businesses outsource work to narrow the wage distribution of their employees, or to avoid paying efficiency wages or rents to workers in occupations with low wages in the labor market.

Appendices B-D describe the relationship between occupational homogeneity and wages when imputed data are not included, when defining employers by Employer Tax Identification Number (EIN) rather than establishments, and separately for states with high and low unionization rates.

This section has described the relationship observed between occupational homogeneity and wages. It cannot say whether employer homogeneity "causes" lower wages for workers in lower-wage occupations. The data used in this paper do not allow me to measure whether differences in unmeasured skills and tasks—within the same occupation—might explain some of the difference in wages between workers in more and less homogenous workplaces. For example, janitors who work in the janitorial services industry may lack some specialized skills of janitors in other industries and may perform somewhat different tasks than those employed in other industries. However, the many U.S. examples described in Weil (2014) and the labor force histories of German workers whose jobs are outsourced, as documented in Goldschmidt and Schmieder (2015) provide evidence that some portion of the observed relationship between employer homogeneity and wages is causal. The estimates in thus section should thus be considered an upper bound for the size of the causal impact of employer homogeneity on wages.

IV: Trends in Occupational Homogeneity Measures

Understanding trends in occupational homogeneity measures is complicated by contemporaneous changes in the overall occupational composition of the labor force. As described by Autor, Katz, and Kearney (2006, 2008), among others, employment in typically low-wage and typically high-wage occupations has increased, while employment in many typically middle wage-occupations has decreased. Figure 1 shows employment over time for occupational quintiles in the OEWS. Employment polarization is clear in the OEWS data: there is an increasing fraction of employment over time in the top quintile, with a decreasing fraction of employment in the middle quintile. This polarization means that if we entirely ignore the grouping of employment into establishments and if occupation-level wages stay constant, the portion of the variance of ln(wages) for all workers due to wage variation between occupations will mechanically increase (from .201 in 2004 to .224 in 2019). There is little mechanical relationship between overall changes in employment by occupation and the Herfindahl-Hirschman index: a version of the Herfindahl-Hirschman index that pools workers across all employers varies only between .0277 and .0283 over this period, with a slightly decreasing time trend.

However, the polarization of employment is not happening evenly across employers. Figure 2 uses the same five quintiles of occupations by typical wages used in Table 2 and shows the fraction of workers in each quintile of occupations who work in establishments with only workers in occupation in their own quintile of the wage distribution. It is unsurprising that workers in the highest-paid quintile of occupations are increasingly likely to work only with this growing quintile of occupations. However, Figure 2 shows that workers in the middle quintiles of occupations, with flat or declining employment, are also increasingly likely over time to have coworkers in occupations with similar wages.

Further descriptive graphical evidence on the polarization of employment across employers is presented in Figure 3. This figure shows how the polarization of employment by quintile is happening by employer size. Weil (2014) describes how large corporations have shed many low-wage tasks by outsourcing them to other companies, which repeatedly subcontract them to smaller and smaller employers. Figure 3 shows that establishment size plays a role in the increasing segregation of workers in the lowest-paid quintile of occupations and workers in the highest-paid quintile of occupations into separate establishments, following the pattern Weil describes. Rising shares of employment for the lowest-paid quintile of occupations occurred only in establishments of less than 100 workers, while rising share of employment for the highest-paid quintile of occupations occurred more sharply in establishments of 100 or more workers.[8] Since Weil describes the outsourcing of low-wage workers to small, homogenous employers, I examine trends in the percentage of workers in the lowest-paid quintile of occupations who work in establishments of less than 100 employees with below-median levels of the partial predicted variance of ln(wages). This percentage increases from 67% in 2004 to 71% in 2010, before falling to 69% in 2016 and 68% in 2019.

Figure 4 displays time trends for the overall measures of occupational homogeneity. This figure includes mean values of occupational homogeneity at the establishment level for each reference date and date-specific coefficients from regressions

(4) $OccHomogeneity_{ijt} = \alpha_t Reference\ Date_t + \beta X_{ijt} + \varepsilon_{ijt}$,

where X_{ijt} are other observable characteristics of individual *i* (occupation) and employer *j* (industry, geography, and size) at time *t*. The same figure also shows (on a second vertical axis) counterfactual levels of occupational homogeneity at the national level, ignoring the grouping of employment into employers. These counterfactuals show the changes in occupational homogeneity due to changes in the overall occupational distribution. Figure 4 shows that both raw and regression-adjusted levels of occupational homogeneity have increased by the normalized HHI measure, despite the small decrease in this measure expected from changes in the occupational distribution. It also shows very little change in the raw or regression-adjusted levels of occupational homogeneity by the partial predicated variance of ln(wages) measure, despite the substantial increase in this measure expected from changes in the occupational distribution.

---

[8] Patterns are similar for establishments of 1-49 workers and establishments of 50-99 workers. Patterns are also quite similar when using EIN size instead of establishment size.

These trends are further documented (overall, and for each quintile of the occupational distribution) with regression results in Table 3. The regressions in this Table have the form

(5) $OccHomogeneity_{ijt} = \alpha_t Decades_t + \beta X_{ijt} + \varepsilon_{ijt}$

where Decades is a continuous measure of time in decades since 2004, and X_{ijt} are other observable characteristics of individual *i* (occupation) and employer *j* (industry, geography, and size) at time *t*. Trend regression results for equation (4) are shown in Table 3. The first two rows of Table 3 show an increase over time in the normalized Herfindahl-Hirschman measure of the occupational homogeneity of employers overall, but changes in occupations and employer characteristics explain about 75% of this increase. The partial predicted variance of ln(wages) measure of occupational homogeneity shows no overall change over time, although there is a very small increase in this measure (a decrease in occupational homogeneity over time) after accounting for changes in occupations and employer characteristics.

The subgroup rows of Table 3 show the greatest increases over time in the normalized Herfindahl-Hirschman measure of occupational homogeneity—after including controls for occupation and establishment characteristics—occur in the bottom and top quintiles of occupations. For the partial predicted variance measure of occupational homogeneity, the pattern is similar for the highest paid quintile of occupations. Only for the highest-paid quintile of occupations is there a consistent decrease in this measure over time (indicating increased occupational homogeneity)—after including controls for occupation and establishment characteristics. In results not shown, the partial predicted variance measure of occupational homogeneity declined sharply for workers in the lowest-paid quintile of workers from 2004 to

2007—both with and without controlling for occupation and establishment characteristics—but this trend reversed after 2010.

These findings are not consistent with the idea that increasing wage inequality in the United States during this period led to increased overall employer occupational homogeneity by specializing in employing either low-wage or high-wage occupations (at least not after 2010). However, as noted above, Dey, Handwerker, Piccone, and Voorheis document a substantial decline in wage inequality in the United States in the later part of this period. Insofar as the intersection of wage inequality growth with the business environment increased incentives for employers to specialize in either low-wage or high-wage work, declining wage inequality should reduce these same incentives for specialization by wage groups. Declining wage inequality would not affect other reasons to outsource, such as the ability to smooth workload and the economies of scale that are possible for providers of specialized services, and changes in occupational homogeneity for these reasons will be captured by the Herfindahl-Hirschman measure of the occupational homogeneity of employers.

Appendices B-D describe employer occupational homogeneity trends when imputed data are not included, when defining employers by Employer Tax Identification Number (EIN) rather than establishments, and separately for states with high and low unionization rates.

V. Occupational Homogeneity and Wage Inequality

The association between occupational homogeneity and lower wages—particularly for workers in lower-wage occupations—coupled with some evidence of growing occupational homogeneity of employers, suggests a role for occupational homogeneity in explaining wage inequality. Barth, Bryson, Davis, and Freeman (2016) highlighted that most inequality growth from 1977-2009 was between establishments, and is not explained by industry or geography. Moreover, Song, Price, Guvenen, and von Wachter (2019) show that the vast majority of pay-inequality growth at small and medium-sized firms in the United States from 1978-2013 was due to increasing segregation and sorting of workers who earn higher pay—without describing what about these workers makes them higher-paid workers—to firms that pay higher wages. Weil (2014) speculated that increased fissuring of employers could exacerbate wage inequality, but he did not have data to measure this directly. This section presents evidence showing that changes in occupational homogeneity contribute to wage inequality and to wage inequality between employers.

I use Rios-Avila's implementation of Fortin, Firpo, and Lemieux's Recentered Influence Function (RIF) Decomposition method to decompose changes in real ln(wage) variance from the 2004 reference date to the 2016 reference date[9] into portions that can be explained by the changing composition of workers by occupation, and the changing composition of their employing establishments by industry, geography, size, and occupational homogeneity. Because

[9] The 2016 reference date is chosen as the end date for this analysis because the OEWS data show a sharp contraction in wage variance from 2016 to 2019, and so there is no overall wage variance growth to explain over the full 2004 to 2019 period. For a more extensive discussion of the fall in wage variance in OEWS data during this period, see Dey, Piccone, Handwerker, and Voorheis.

the occupational homogeneity measures are continuous rather than categorical variables, these variables are divided into quartiles for this reweighting exercise. The evidence in Table 3 shows that occupational homogeneity is changing in different ways for different quintiles of occupations. Thus, I interact occupational homogeneity variables with the same quintiles of occupation used above.[10] In addition, I add a dummy variable for lowest-wage quintile occupations employed in establishments of less than 100 workers that are in the bottom half of the predicted variance distribution to the vector of indicator variables describing the predicted variance measure of occupational homogeneity.

Results are shown in Table 4. The changing composition of employment by industry, geography, establishment sizes, occupational quintiles, and the categories of occupational homogeneity described above can more than explain all of the growth in ln(wage) variance from 2004 to 2016. Decomposing the change in ln(wage) variance by source, by far the category which most explains wage inequality growth is the changing pattern of employment by occupational quintiles (employment polarization), which explains 79% of wage inequality growth. Changes in employment by the occupational homogeneity of employers explains 12% of wage inequality growth—7% for changing values of the Hirschman-Herfindahl index, and 5% for the changing values of predicted employer wage variance based on between-occupation wage variation, with much of this coming from the dummy variable for low wage occupations employed in small establishments with low predicted ln(wage) variance. Changing industry composition explains much of the remaining growth in wage variation.

[10] This follows the example of Goldschmidt and Schmieder (section V.C.), who use indicators for deciles of the firm wage effect interacted with dummies for frequently outsourced occupations.

To examine the impact of changing occupational homogeneity on the growth of wage variance between establishments, I use the Dinardo-Fortin-Lemieux (DFL) 1996 method. This method calculates counterfactual wage distributions by reweighting observable characteristics in the later period (2016) to their distributions in the earlier period (2004). The overall variance of real ln(wages) increased from 0.362 for the 2004 reference date to 0.380 for the 2016 reference date, and most of this increase is due to between-establishment wage variance increasing from 0.205 to 0.220. Reweighting the 2016 data to the 2004 distribution of employment by quartiles of both occupational homogeneity measures (without interacting these occupational homogeneity measures with occupational quintiles, to avoid also capturing the impact of employment polarization by occupation) and the indicator for workers in typically lower-wage occupations employed in small homogenous establishments, the between-establishment wage variance would be .213 rather than the actual .220. This reweighting explains about half of the wage variance growth between establishments.

Wage variation, including the between-establishments portion of wage variation, declined from 2016 to 2019 (with the between-establishments portion of ln(wage) variance falling from .220 to .208). However, applying this reweighing method, the between establishments portion of ln (wage) variance in 2019 would have been still lower (.203) under the 2004 distribution of occupational homogeneity variables.[11]

[11] Reweighting 2016 or 2019 data to the 2004 distributions of occupational homogeneity variables without interacting these occupational homogeneity variables with occupations does not fully capture the impact of changes in occupational homogeneity, but has the advantage of not being co-mingled with changes in employment by occupation. This reweighting reduces the ln(wage) variance between establishments in the 2016 and 2019 data without reducing overall ln(wage) variance. Reweighting 2016 or 2019 data to the 2004 distributions of occupational homogeneity variables and also to the 2004 distributions of the interactions of occupational homogeneity with occupational quintiles more completely captures the impact of changing occupational homogeneity in different parts of the wage distribution, but is co-mingled with changes in the occupational distribution from 2004 to 2016. This reduces ln(wage) variance between establishments even further (from .220 to .197 in 2016 and from .208 to .184 in 2019) and also reduces the overall ln(wage) variance (from .380 to .365 in 2016 and from .360 to .339 in 2019).

In sum, these results show that changes in occupational homogeneity are an important part of growing wage inequality for the lower 98.5% of the wage distribution. Both the Herfindahl-Hirschman measures of overall occupational homogeneity and the growth of employment in low-wage occupations at small employers with below-median predicted variances of wages (employing few typically high wage occupations) are important for wage inequality growth during this period.

VI. Summary: Outsourcing and increasing wage inequality

While many authors have studied the growth in wage inequality between employers and others have studied the impact of outsourcing on wages for specific occupations, this paper is among the first to study the empirical relationship between the changing distribution of occupations between employers and changing wage inequality in the United States. It focuses on the occupational homogeneity of employers as a measure of outsourcing, offering an economy-wide view of how occupational homogeneity impacts wages and wage inequality.

Consistent with previous literature on outsourcing, there is greater occupational homogeneity for the occupations used in case studies of outsourcing by Abraham and Taylor (1996); Dube and Kaplan (2010); Dey, Houseman, and Polivka (2010); and Goldschmidt and Schmeider (2015) when these occupations are employed in establishments in the industry to which they are outsourced. For example, employer occupational homogeneity is higher for janitors when they are employed in establishments in the janitorial services industry than when they are employed in other industries.

The advantage of using occupational homogeneity to measure outsourcing is that these occupational homogeneity measures can be calculated for every employee of every employer, not only for “case study” occupations and industries. This paper shows that economy-wide, employer occupational homogeneity is related to wage levels. It has a particularly strong negative wage association for workers in occupations that are typically low paid, even after controlling for the occupations of employees and various observable characteristics of their employers. In contrast, workers in the highest paid quintile of occupations are paid more if they have fewer co-workers in typically low-wage occupations, after controlling for their own occupations and the observable characteristics of their employers.

Coarse measures of occupational homogeneity show that the changing composition of the workforce is not occurring evenly across employers. Low-wage occupations are growing in smaller employers and shrinking in larger employers, while the growth of high-wage occupations is concentrated in large employers. Workers in every part of the wage distribution have an increasing share over time of coworkers in their own part of the wage distribution.

Employer occupational homogeneity as measured by the Herfindahl-Hirschman index of occupations within employers shows steady increases in employer homogeneity over time, particularly for workers in relatively low-paid occupations. These patterns are consistent with the idea that companies are “de-verticalizing” by outsourcing functions not part of their “core competencies.”

Occupational homogeneity trends in the predicted wage variance measure based on the occupational composition of employers are affected by contemporaneous changes in the overall occupational composition of the workforce. Yet this measure of occupational homogeneity shows increased homogeneity over time for workers in typically high-wage occupations overall, and for workers in typically low-wage occupations in the early part of the period, when wage inequality was increasing. By this measure, occupational homogeneity was very slightly declining overall between 2004 and 2019. This suggests no overall increase throughout this period in the outsourcing of specifically low-wage occupations. However, employer incentives for this form of outsourcing increase when wage inequality increases, and in the latter part of this period, these data show wage inequality declining. This decline in wage inequality would reduce incentives for outsourcing for the purpose of narrowing the wage distribution of employees.

Song, Price, Guvenen, and von Wachter (2019) show that the vast majority of pay-inequality growth at small and medium-sized firms is due to the increasing segregation and sorting of workers who earn lower pay—without describing what about these workers makes them lower-paid workers—to firms that pay lower wages. To the extent that workers stay in the same occupations over time, occupation is exactly the sort of characteristic that would lead workers to earn different pay levels than other workers, no matter their employer. This paper provides evidence that the sorting and segregation of high and low-wage workers happens at the occupational level, contributing to the growth of between-employer wage inequality.

Although the data used in this paper cannot show changes in the wage distribution for the very highest 1.3% of wage-earners, they are well suited to measure the contribution of employers' occupational homogeneity to wage inequality growth for the remaining 98.7% of the wage distribution. Decompositions of ln(wage) variance growth in these data show the growing polarization of employment can explain the majority of inequality growth between 2004 and 2016, and the changing distribution of occupational homogeneity can explain much of the remainder. Although wage inequality in these data fell sharply after 2016, the growing separation of workers doing different types of work was an important component of the wage inequality growth observed until 2016, and the wage convergence observed from 2016 to 2019 would have been even greater in the absence of these changes in occupational homogeneity.

REFERENCES

Abowd, John M., Francis Kramarz, and David Margolis (1999). “High Wage Workers and High Wage Firms.” *Econometrica*, pp. 251-334.

Abraham, Katharine G. and James R. Spletzer (2010). “Are the New Jobs Good Jobs?,” in *Labor in the New Economy*, edited by Katharine G. Abraham, James R. Spletzer, and Michael Harper. University of Chicago Press.

Abraham, Katherine G and Susan K. Taylor (1996). “Firms’ use of Outside Contractors: Theory and Evidence.” *Journal of Labor Economics*, 14(3): 394-424.

Acemoglu, Daron (2002). “Technical Change, Inequality, and the Labor Market.” *Journal of Economic Literature*, Vol. 40, No. 1, pp. 7-72.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2006). “The Polarization of the U.S. Labor Market,” *American Economic Association Annual Meeting Papers and Proceedings*, pp. 189-194.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2008). “Trends in U.S. Wage Inequality: Re-Assessing the Revisionists.” *Review of Economics and Statistics*, Vol. 90, No. 2, pp. 300-323.

Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman (2016). “It's Where You Work: Increases in Earnings Dispersion across Establishments and Individuals in the U.S.” *Journal of Labor Economics*, 34(S2): S67-S97.

Bilal, Adrien, and Hugo Lhuillier. Outsourcing, Inequality and Aggregate Output. No. w29348. National Bureau of Economic Research, 2021.

Bound, John and George Johnson (1992). “Changes in the Structure of Wages in the 1980's: An Evaluation of Alternative Explanations.” *The American Economic Review*, Vol. 82, No. 3, pp. 371-392.

Bronars, Stephen G. and Melissa Famulari (1997). “Wage, Tenure, and Wage Growth Variation Within and Across Establishments.” *Journal of Labor Economics*, pp. 285-317.

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.

Dey, Matthew, David Piccone Jr., Elizabeth Weber Handwerker, and John Voorheis (2022). “Were Wages Conversing During the 2010s Expansion?” *Monthly Labor Review*.

Dey, Matthew, and Elizabeth Weber Handwerker (2016). "Longitudinal data from the Occupational Employment and Wage Statistics survey." *Monthly Labor Review*.

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 J. Harper, eds. Chicago: University of Chicago Press, 2010, pp. 267-304

Davis, Steve J. and John Haltiwanger (1991). "Wage Dispersion Between and Within U.S. Manufacturing Plants." *Brookings Papers on Economic Activity*, pp. 115-200.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, Vol. 64, No. 5, pp. 1001-1044.

Drenik, Andres, Simon Jäger, Miguel Pascuel Plotkin, and Benjamin Schoefer (2021). Paying Outsourced Labor: Direct Evidence from Linked Temp Agency-Worker-Client Data. *The Review of Economics and Statistics*, pp. 1-28.

Dube, Arindrajit, and Ethan Kaplan (2010). "DOES outsourcing reduce wages in the low-wage service occupations? Evidence from janitors and guards." *ILR Review* 63.2: 287-306.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske (2004). "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment." *Journal of Labor Economics*, Vol. 22, No. 2, pp. 397-429.

Employee Benefit Research Institute (EBRI) (2009). *Fundamentals of Employee Benefit Programs, 6th ed.* http://www.ebri.org/publications/books/index.cfm?fa=fundamentals

Erickcek, George, Susan Houseman, and Arne Kalleberg (2003). "The Effects of Temporary Services and Contracting Out on Low-Skilled Workers: Evidence from Auto Suppliers, Hospitals, and Public Schools," in *Low-Wage America: How Employers are Reshaping Opportunity in the Workplace*, edited by Eileen Appelbaum, Annette Bernhardt, and Richard J. Murnane. New York: Russell Sage Foundation, pp.368-403.

Firpo, Sergio P., Nicole M. Fortin, and Thomas Lemieux (2018). "Decomposing Wage Distributions Using Recentered Influence Function Regressions." *Econometrics* 6 (3):41.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo (2011). "Decomposition Methods in Economics," *Handbook of labor economics* 4 (2011): 1-102.

Goldschmidt, Deborah and Johannes F. Schmieder (2017). "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure," *The Quarterly Journal of Economics* 132(3):1165-1217.

Groshen, Erica L. (1991a). "Sources of Intra-Industry Wage Dispersion: How Much do Employers Matter?" *The Quarterly Journal of Economics*, Vol. 106, No. 3, pp. 869-884.

Groshen, Erica L. (1991b). “Five Reasons Why Wages Vary Among Employers.” *Industrial Relations*, Vol. 30, No. 3, pp. 350-381.

Handwerker, Elizabeth Weber, & Mason, Lowell G. (2013). Linking firms with establishments in BLS microdata. *Monthly Labor Review*, Vol 136, No 14.

Handwerker, Elizabeth Weber and James R. Spletzer (2014). “Measuring the distribution of wages in the United States from 1996-2010 with the Occupational Employment Survey” *Monthly Labor Review*, May 2014

Handwerker, Elizabeth Weber, and James R. Spletzer (2016). "The Role of Establishments and the Concentration of Occupations in Wage Inequality." *Inequality: Causes and Consequences.* Emerald Group Publishing Limited, 2016. 167-193.

Juhn, Chinhui, Kevin M. Murphy, Brooks Pierce (1993). “Wage Inequality and the Rise in Returns to Skill Wage Inequality and the Rise in Returns to Skill.” *The Journal of Political Economy*, Vol. 101, No. 3, pp. 410-442.

Katz, Lawrence F. and Autor, David H. (1999). “Changes in the Wage Structure and Earnings Inequality.” In *Handbook of Labor Economics* (volume 3A), edited by Orley Ashenfelter and David Card, Elsevier, pp. 1463-1555.

Katz, Lawrence F. and Kevin M. Murphy (1992). “Changes in Relative Wages, 1963-1987: Supply and Demand Factors.” *The Quarterly Journal of Economics*, Vol. 107, No. 1, pp. 35-78.

Lane, Julia. I., Salmon, Laurie A., & Spletzer, James R. (2007). Establishment wage differentials. *Monthly Labor Review*, Vol. 130, No.3.

Lemieux, Thomas (2006). “Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?” *American Economic Review*, Vol. 96, No. 3, pp. 461-498.

Perun, Pamela, "FAQs about Employees and Employee Benefits" (2010). *Georgetown Law School Memos and Fact Sheets*. Paper 46. Available at http://scholarship.law.georgetown.edu/legal/46

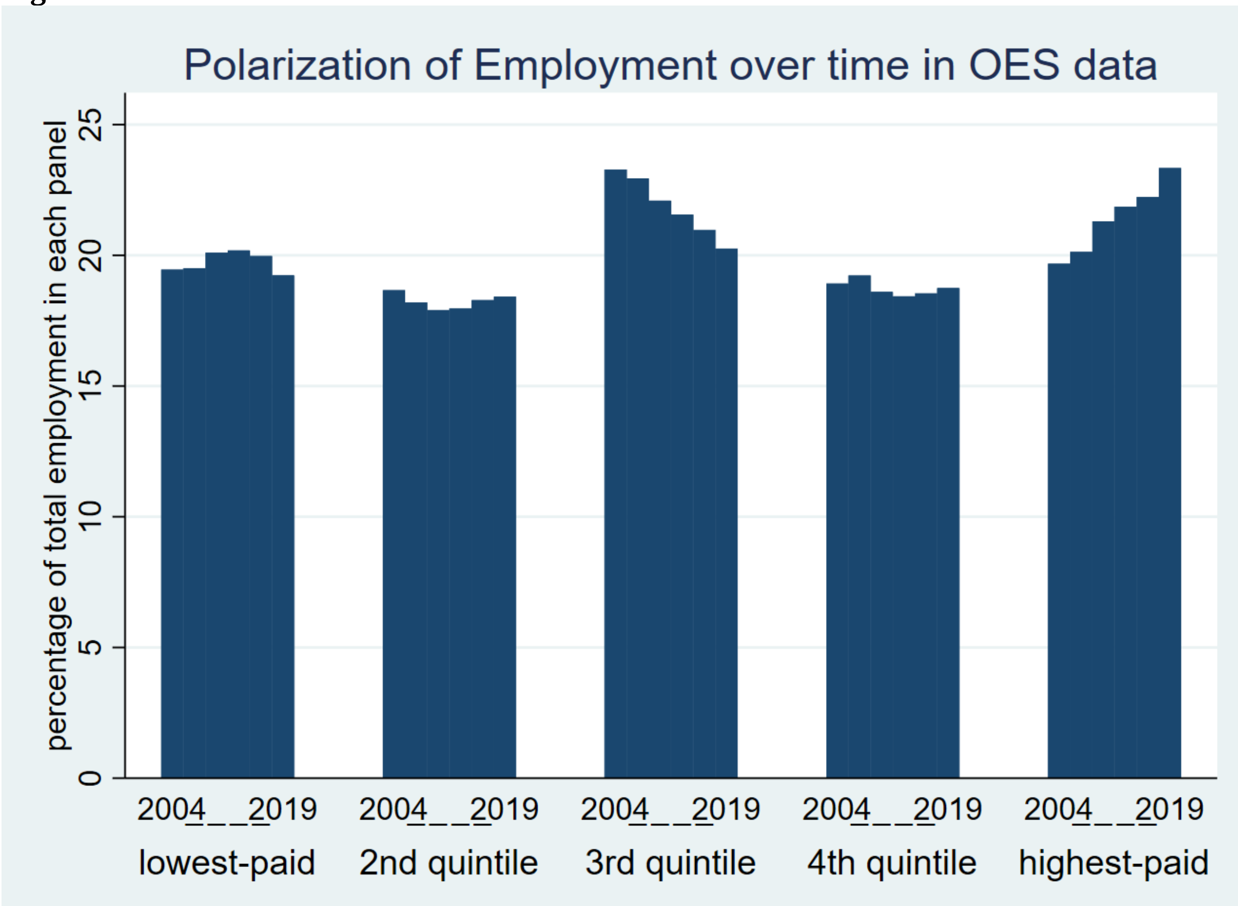
Rios-Avila, Fernando (2020). “Recentered Influence Functions (RIF) in Stata: RIF Regression and RIF Decomposition” *The Stata Journal.* Vol 20, No. 1, pp. 51-94.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. (2019). "Firming up inequality." *The Quarterly Journal of Economics* Vol. 134, No 1, pp. 1-50.

United States Bureau of Labor Statistics (2021). BLS handbook of methods. https://www.bls.gov/opub/hom/oews/

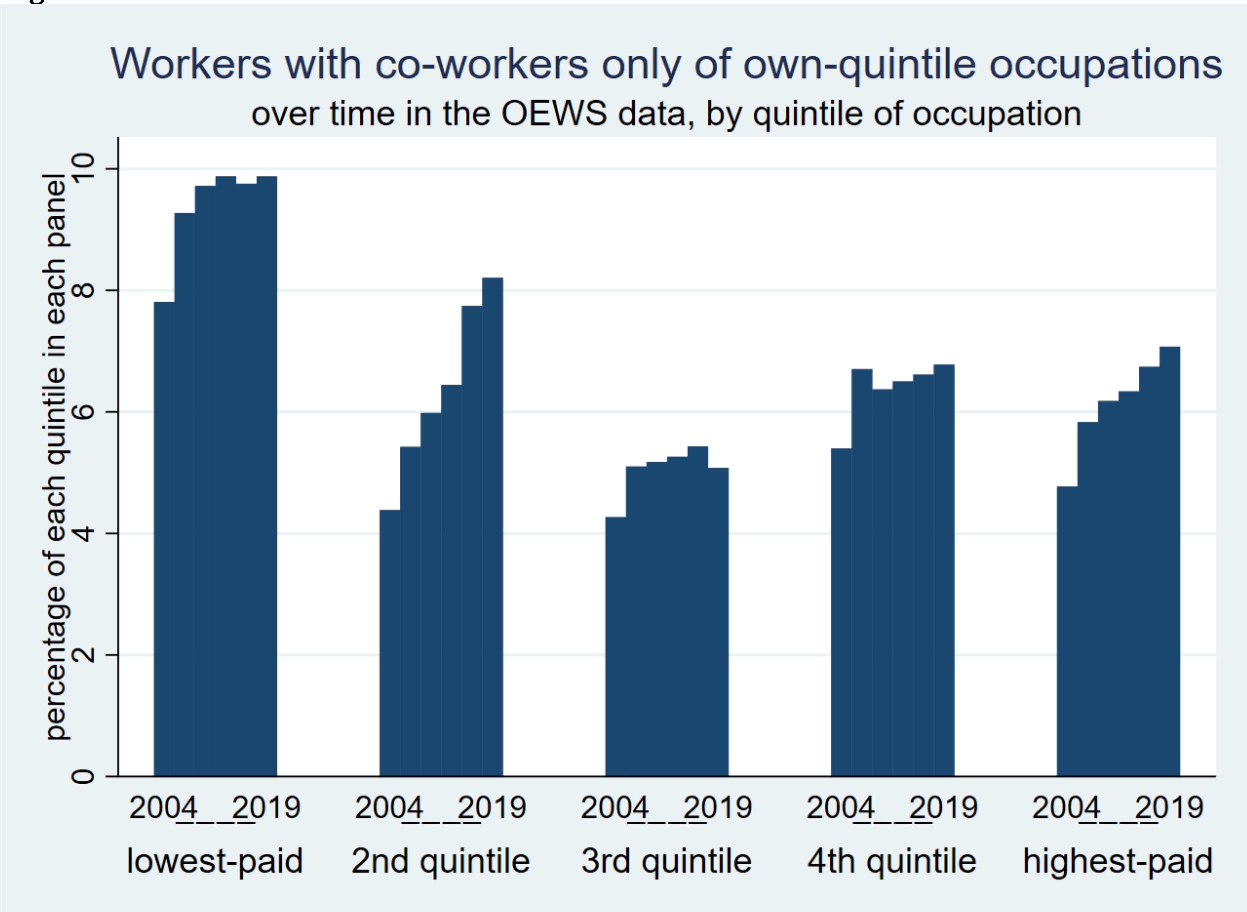
Weil, David. The fissured workplace. Harvard University Press, 2014.

Figure 1:



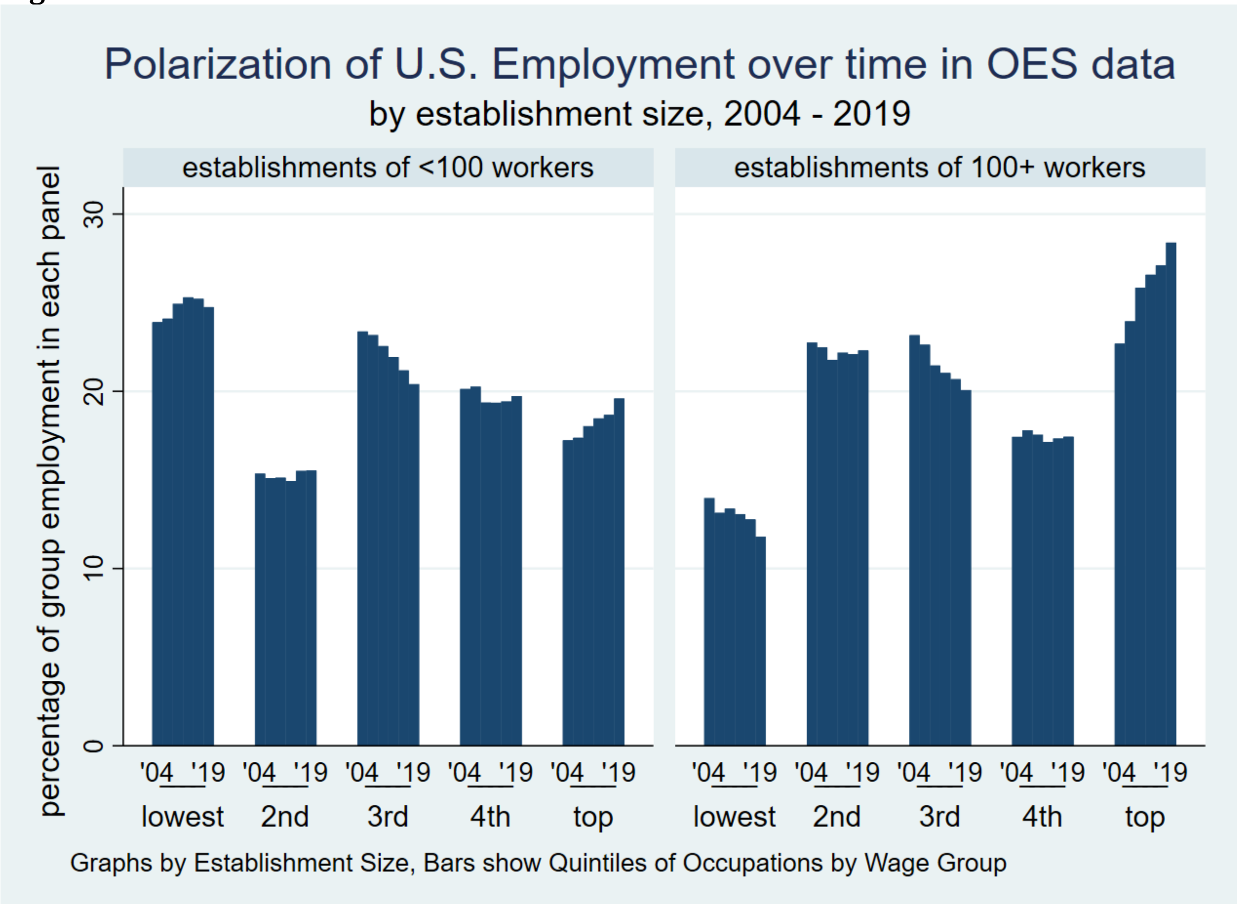
Note: The 94,626,298employer-occupation-wage level observations from 6 panels for each reference date are used to calculate overall average wage levels and employment levels in 2004, 2007, 2010, 2013, 2016, and 2019. These are grouped into quintiles of minor occupational groups (3-digit SOC groups) by average wage levels (as shown in Appendix A). Quintiles may have slightly more or less than 20% of employment because of large occupational groups. This figure shows the percentage of employment in each occupational quintile in each panel of OEWS data, from November 2004 (collected from 2001 to 2004) through November 2019 (collected from 2017 to 2019).

Figure 2:



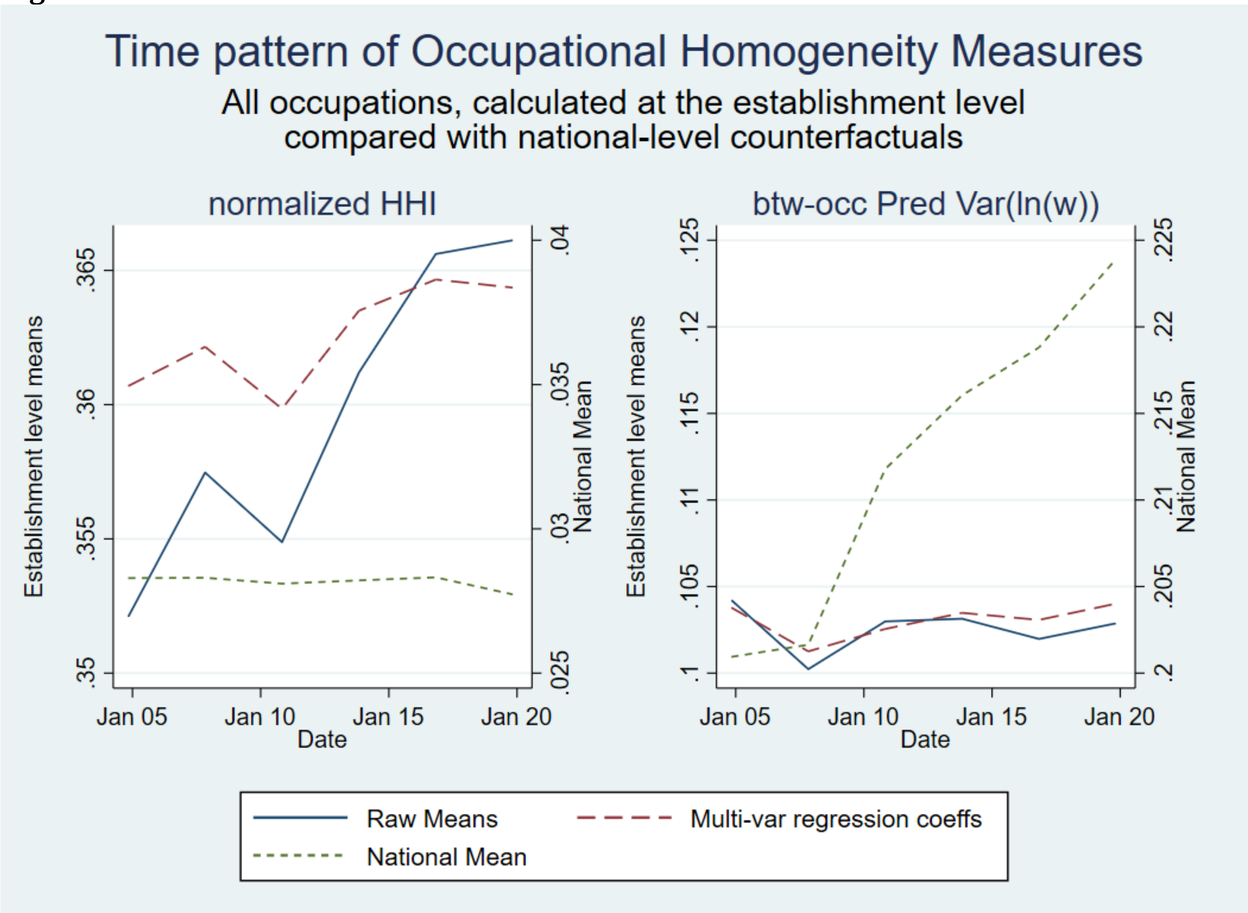
Note: The 94,626,298employer-occupation-wage level observations from 6 panels for each reference date are used to calculate overall average wage levels and employment levels in 2004, 2007, 2010, 2013, 2016, and 2019. These are grouped into quintiles of occupation by average occupational wage levels (as shown in Appendix A). This figure shows the percentage of workers in each quintile employed in establishments that have only workers in their own quintile, by panel (from November 2004 through November 2019). For example, the subgraph at the top left shows the fraction of workers in the lowest-quintile of occupations who have no co-workers in each other quintile of occupations, for each panel of the OEWS data.

Figure 3:



Note: The 94,626,298 employer-occupation-wage level observations from 6 panels for each reference date are used to calculate overall average wage levels and employment levels in 2004, 2007, 2010, 2013, 2016, and 2019. These are grouped into quintiles of minor occupational groups (3-digit SOC groups) by average wage levels (as shown in Appendix A). Quintiles may have slightly more or less than 20% of employment because of large occupational groups. This figure shows the percentage of employment in each establishment size group in each occupational quintile in each panel of OEWS data, from November 2004 (collected from 2001 to 2004) through November 2019 (collected from 2017 to 2019).

Figure 4:



Note: Raw Means show average measured levels of occupational homogeneity at the establishment level for the 94,626,298 employer-occupation-wage level observations from 6 panels for each reference date. Normalized Herfindahl indices of Occupational Homogeneity for establishments are calculated at the minor occupational group level and are normalized for the number of employees in the establishment. The Partial Predicted Variances of ln(wages) for establishments are based on employment by the same minor occupational groups within establishments, and do not include predicted within-occupational group wage variation. Multi-variate regression coefficients are the coefficients α from regressions of the form $Occupation\ Homogeneity_{ijt} = \alpha Reference\ Date_t + \beta X_{ijt} + \varepsilon_{ijt}$, where X_{ijt} includes occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, establishment size class, continuous establishment size, and state of location. Regressions are at the establishment-occupation-wage interval level, weighted by employment. National Means show values of Occupational Homogeneity based on the national occupational distribution at each reference date, ignoring the grouping of occupations into establishments. These have much lower levels of Occupational Homogeneity than the establishment level means of the same variables, and so their values are plotted on the right vertical axis. For each occupational homogeneity variable, both vertical axes use the same scale and a different range of values.

Table 1: Mean Values of Occupational Homogeneity for Specified Occupations and Industries, 2004-2019

Occupation and Industry	Avg ln(wage)	Mean Value of Occupational Homogeneity: Normalized Herfindahl of Occupational Homogeneity for the establishment	Mean Value of Occupational Homogeneity: Partial Predicted Variance of Wages for the establishment
Food preparation and serving (SOC 35)			
within Food Services (NAICS 722) – 81%	2.02	.464	.056
within all other industries – 19%	2.12	.237	.122
Janitors (SOC 372011)			
within Janitorial Services (NAICS 561720) –47%	2.09	.824	.043
within all other industries –53%	2.17	.286	.117
Security Guards (SOC 339032)			
within Security Guard Srvcs (NAICS 561612) –61%	2.20	.871	.029
within all other industries –39%	2.32	.314	.117
Truck Drivers (SOC 53303)			
within Truck Transportation (NAICS 484) –30%	2.68	.593	.038
within all other industries –70%	2.46	.339	.083
Accountants (SOC 132011)			
within Accounting Services (NAICS 541211) –25%	3.22	.485	.080
within all other industries –75%	3.16	.223	.132
Computer Occupations (SOC 151)			
within Computer Services (NAICS 5415) –28%	3.34	.500	.056
within all other industries –72%	3.29	.279	.115
Engineers (SOC 172)			
within Engineering Services (NAICS 54133) –21%	3.41	.320	.091
within all other industries –79%	3.43	.226	.124
Lawyers (SOC 231011)			
within Law Offices (NAICS 54111) –81%	3.76	.283	.277
within all other industries –19%	3.85	.227	.152
The most dispersed quintile of occupational groups			
within Business Services (NAICS 561) –9%	2.47	.435	.076
within all other industries -91%	2.75	.272	.122

Notes: Data is pooled across 94,928,505 employer-occupation-wage level observations from 6 panels of data for each reference date from November 2004 through November 2019. Normalized Herfindahl indices of Occupational Homogeneity for the establishment are calculated at the minor occupational group level and are normalized for the number of employees in the establishment. Partial Predicted Variances of Wages for the establishment are based on employment by minor occupational group within the establishment, and do not include predicted within-occupational group wage variation. The “most dispersed quintile of occupational groups” are the minor occupational groups with the highest HHI indices of employment across industries, and include such groups as Other Office and Administrative Support Workers (SOC 439), Information and Record Clerks (SOC 434), and Other Production Occupations (SOC 519).

Table 2: Regressions of log wages on measures of Occupational Homogeneity

Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment
All Occupations (94,628,505 observations)		
With only date fixed effects	-0.532	1.823
	(0.002)	(0.006)
All Controls	-0.055	0.073
	(0.001)	(0.003)
Lowest-paid quintile of occupations (7,725,178 observations)		
With only date fixed effects	-0.171	0.635
	(0.001)	(0.005)
All Controls	-0.073	0.165
	(0.001)	(0.005)
Second quintile of occupations (11,784,546 observations)		
With only date fixed effects	-0.204	0.690
	(0.002)	(0.006)
All Controls	-0.078	0.346
	(0.002)	(0.005)
Middle quintile of occupations (22,319,551 observations)		
With only date fixed effects	-0.131	0.523
	(0.002)	(0.005)
All Controls	-0.070	0.301
	(0.002)	(0.005)
Fourth quintile of occupations (20,693,841 observations)		
With only date fixed effects	-0.188	0.514
	(0.002)	(0.006)
All Controls	-0.044	0.144
	(0.002)	(0.006)
Highest-paid quintile of occupations (32,105,389 observations)		
With only date fixed effects	-0.106	0.102
	(0.004)	(0.010)
All Controls	-0.014	-0.464
	(0.003)	(0.008)

Notes: These are coefficients α from regressions of the form $\ln(wage_{ijt}) = \alpha OccHomogeneity_{jt} + \beta X_{ijt} + \varepsilon_{ijt}$, where X includes reference date fixed effects, occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, state fixed effects, and establishment size (using fixed effects for establishment size classes as well as continuous establishment size). All coefficients are significant at $p < 0.001$. Regressions are at the establishment-occupation-wage interval level, weighted by employment. Occupations (at the minor occupational group level) found in each quintile are listed in Appendix A. Normalized Herfindahl indices of Occupational Homogeneity for the establishment are calculated at the minor occupational group level, and are normalized for the number of employees in the establishment. Partial Predicted Variances of Wages for the establishment are based on employment by minor occupational group within the establishment, and do not include predicted within-occupational group wage variation. Clustered standard errors in parentheses.

Table 3: Change in mean values of Occupational Homogeneity over time

Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment
All Occupations (94,628,505 observations)		
Raw Trend	0.0096	-0.0000
	(0.0005)	(0.0002)
All Controls	0.0029	0.0008
	(0.0003)	(0.0001)
Lowest-paid quintile of occupations (7,725,178 observations)		
Raw Trend	0.0112	0.0004
	(0.0010)	(0.0002)
All Controls	0.0049	0.0026
	(0.0006)	(0.0002)
Second quintile of occupations (11,784,546 observations)		
Raw Trend	0.0330	-0.0030
	(0.0011)	(0.0002)
All Controls	0.0024	0.0021
	(0.0007)	(0.0002)
Middle quintile of occupations (22,319,551 observations)		
Raw Trend	0.0090	0.0012
	(0.0007)	(0.0002)
All Controls	-0.0003	0.0029
	(0.0005)	(0.0001)
Fourth quintile of occupations (20,693,841 observations)		
Raw Trend	-0.0035	0.0026
	(0.0008)	(0.0002)
All Controls	-0.0000	0.0014
	(0.0005)	(0.0001)
Highest-paid quintile of occupations (32,105,389 observations)		
Raw Trend	0.0051	-0.0054
	(0.0008)	(0.0003)
All Controls	0.0043	-0.0043
	(0.0006)	(0.0002)

Note: These are coefficients α from regressions of the form $Occupation\ Homogeneity_{ijt} = \alpha Decades_t + \beta X_{ijt} + \varepsilon_{ijt}$, where Decades measures time in decades since 2004 and X_{ijt} includes occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, establishment size class, continuous establishment size, and state of location. Regressions are at the establishment-occupation-wage interval level, weighted by employment. Occupations (at the minor occupational group level) found in each quintile are listed in Appendix A. Normalized Herfindahl indices of Occupational Homogeneity for the establishment are calculated at the minor occupational group level, and are normalized for the number of employees in the establishment. Partial Predicted Variances of Wages for the establishment are based on employment by minor occupational group within the establishment, and do not include predicted within-occupational group wage variation. Clustered standard errors are in parentheses.

Table 4: Decomposition of Changes in real ln(wage) variance from 2004 to 2016

Observations in late period (2016): 14,016,725
Observations in early period (2004): 13,012,513

Real log wage variance	Coeff.	Percent	Bootstrapped Standard Deviations
Overall Variance			
Late period (2016 reference date)	0.3803		.0003
Counterfactual variance	0.3591		.0003
Early period (2004 reference date)	0.3621		.0004
Total change	0.0182	100%	.0005
of which explained (by compositional change)	0.0212	116%	.0003
of which unexplained (wage structure change)	-0.0030	-16%	.0004
Explained (compositional effect)			
Total	0.0212	100%	.0003
Pure explained	0.0211	100%	.0003
Specification error	0.0000	0%	.0000
Components of the pure explained effect			
Industry sector (2-digit NAICS)	0.0021	10%	.0001
Geography (Census Division)	0.0004	2%	.0000
Establishment size	-0.0008	-4%	.0000
Occupation quintiles (defined in Appendix A)	0.0167	79%	.0002
Normalized Herfindahl measure of establishments	0.0016	7%	.0001
Partial predicted variance of establishment ln(wages)	0.0011	5%	.0001
Unexplained (wage structure changes)			
Total	-0.0030	100%	.0004
Reweighting error	0.0000	1%	.0000
Pure unexplained	-0.0030	99%	.0004

Notes: These are the results of Rios-Avila's implementation of Fortin, Firpo, and Lemiueux's Recentered Influence Function Decomposition. OEWS data with a 2004 reference date was collected from 2001 to 2004; date with a 2016 reference date was collected from 2014 to 2016. Industry here is grouped into 19 supersectors and geography into 7 Census divisions. Establishment size is measured in 9 categories (1-4 employees, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, and 1000+). Quintiles of occupations are defined in Appendix A. Establishment-level normalized Herfindahl-Hirschman indices of occupations are measured with quartiles of the distribution, interacted with occupational quintiles. The predicted variance of ln(wage) for establishments is also divided into quartiles and interacted with occupational quintiles, with an additional dummy variable for low-wage occupations in establishments of less than 100 workers that are in the bottom half of the predicted variance distribution. Standard deviations are the results of bootstrapping the coefficients with 300 replications.

Data Appendix

This paper uses Occupational Employment and Wage Statistics (OEWS) Survey microdata. The OEWS survey is designed to measure occupational employment and wages in the United States by geography and industry, and is the only such survey of its size and scope, covering all establishments in the United States except those in agriculture, private households, and unincorporated self-employed workers without employees. Every year, approximately 400,000 private and local government establishments are asked to report the number of employees in each occupation paid within specific wage intervals: 200,000 establishments each November and another 200,000 each May. As described in Dey and Handwerker, the OEWS uses a complex sample design intended to minimize the variance of wage estimates for each occupation within industries and geographic areas. Thus, establishments in rarer industries and geographic areas, as well as establishments expected to employ occupations with greater variation in wages have relatively larger probabilities of selection and lower estimation weights.

The OEWS survey form (now largely replaced by electronic collection of payroll reports) is a matrix of detailed occupations and wage intervals. Establishments provide job titles and short descriptions of job duties, which are coded into occupations by staff in state labor agencies. Wage intervals on the OEWS survey form are given in both hourly and annual nominal dollars, with annual earnings that are 2080 times the hourly wage rates. To calculate average wages, the OEWS program obtains the mean of each wage interval for each minor occupational group for each reference period from the National Compensation Survey (NCS). These mean wages are then assigned to all employees in that wage interval. To adjust wage estimates collected at different dates within a three-year sample period, the OEWS program uses the BLS Employment Cost Index for each occupational division.

The OEWS sample design uses 3 years, or 6 panels of data collection, to produce detailed published estimates of employment and wages, with employment weights benchmarked to employment at the time of the last panel and adjustments to wages based on the BLS Employment Cost Index so that wages refer to wage levels in that last panel. It is not designed to produce time series estimates of either employment or wages for any individual occupation, in part because of changes over time in occupational definitions. More information about the OEWS program can be found in the BLS Handbook of Methods.

The OEWS has been using the Standard Occupational Classification System since 1999 and had a change of industry classification systems from SIC to NAICS (2002) soon thereafter. Certain SOC and NAICS codes are combined to make groups consistent across the 2007, 2012, and 2017 NAICS revisions and the 2010 and 2018 revisions to the SOC. Data used in this paper begin in November 2001 to avoid inconsistencies of SOC coding in small establishments during the initial years that the OEWS program used this coding system, as described by Abraham and Spletzer. The SOC revisions of 2010 and 2018 are much less substantial and can be addressed by (relatively small) aggregations of occupations. For example, the 2010 revision to the Standard Occupational Classification System split Registered Nurses into four occupations: Registered Nurses, Nurse Anesthetists, Nurse Midwives, and Nurse Practitioners, but for this paper all four are recoded as one occupation throughout.

Handwerker and Spletzer (2016) examine the decomposition of total wage variance in the OEWS into its within-establishment and between establishment components at length. Updating their findings to 2016, from Fall 1999 through November 2016, 60% of wage variance in the OEWS was between establishments, while all growth in overall wage variance over this period was between establishments. From November 2016 to November 2019, overall wage variance in the OEWS fell by 5%, of which 55% of the decline was due falling wage variance between establishments.

Dey, Handwerker, Piccone, and Voorheis document that OEWS data display substantial overall wage convergence from 2013 to 2019, particularly from 2016 to 2019. This wage convergence was due to strong wage growth for workers in low-wage occupations, rather than to changes in the occupational distribution. They show these patterns in the OEWS data are broadly consistent with patterns observed in the CPS and in wage data collected by the income tax system. However, the declines in overall wage inequality in the OEWS from 2016 to 2019 are stronger than those observed in the CPS data because the CPS data show greater wage growth for high-wage workers during this period than is observed in the OEWS.

Data Appendix Table 1: Summary Statistics

Variable	Observations	Employment represented	Weighted Mean	Min	Max	Standard Deviation
OEWS real wage	94,626,298	685,540,362	16.55	5.25	109.91	13.99
OEWS ln(wage)	94,626,298	685,540,362	2.59	1.66	4.70	0.61
Measured var(ln(wage)) of establishments	94,626,298	685,540,362	0.167	0.000	2.222	0.134
Herfindahl-Hirschman index of estab employment (by minor occupational group)	94,626,298	685,540,362	0.400	0.031	1.000	0.248
Normalized Herfindahl-Hirschman index of establishment employment	94,626,298	685,540,362	0.360	0.000	1.000	0.250
Portion of the predicted var of ln(wages) for each establishment due to variation in wages **between** minor occupational groups	94,626,298	685,540,362	0.103	0.000	0.683	0.076
Establishment-level employment	94,626,298	685,540,362	560	1	confidential	2129
Reference date for observation	94,626,298	685,540,362	2011.75	2004	2019	5.17
Decades since 2004	94,626,298	685,540,362	0.78	0.00	1.50	0.52

Variable Distributions	Observations	Employment represented	Fraction of Employment	Establishment observations
Quintiles of occupation – occupations are listed in Appendix A				
Lowest-paid quintile of occupations	7,725,178	135,364,590	19.7%	
2nd quintile of occupations	11,782,345	124,699,305	18.2%	
Middle quintile of occupations	22,319,551	149,537,779	21.8%	
4th quintile of occupations	20,693,840	128,587,303	18.8%	
Highest paid quintile of occupations	32,105,384	147,351,384	21.5%	
Major industry groups (2-digit)				
Agriculture, Forestry, Fishing and Hunting,	190,214	2,392,838	0.3%	30,319
Mining, Quarrying, and Oil and Gas Extraction	723,334	3,922,932	0.6%	42,031
Utilities	745,972	3,313,182	0.5%	36,539
Construction	5,165,396	40,525,625	5.9%	503,814
Manufacturing	15,633,177	76,728,433	11.2%	665,284
Wholesale Trade	6,550,147	34,747,748	5.1%	473,460
Retail Trade	9,858,784	92,921,634	13.6%	765,245
Transportation and Warehousing	2,898,187	28,394,498	4.1%	237,170
Information	3,531,304	17,192,071	2.5%	190,994
Finance and Insurance	5,954,353	38,345,572	5.6%	375,429
Real Estate and Rental and Leasing	1,376,340	9,043,767	1.3%	155,281
Professional, Scientific, and Technical Services	7,869,054	48,579,577	7.1%	587,715
Management of Companies and Enterprises	3,176,001	12,116,694	1.8%	73,799
Administrative and Support and Waste Management and Remediation Services	5,756,520	50,832,245	7.4%	440,223
Educational Services	2,853,391	15,726,201	2.3%	118,543
Health Care and Social Assistance	12,836,692	101,606,031	14.8%	686,792
Arts, Entertainment, and Recreation	2,116,635	12,294,224	1.8%	155,454
Accommodation and Food Services	4,050,790	73,089,407	10.7%	337,523

Variable Distributions	Observations	Employment represented	Fraction of Employment	Establishment observations
Other Services	3,340,007	23,767,682	3.5%	424,434
Major Occupational Groups (2-digit)				
Management Occupations	10,920,055	34,105,570	5.0%	
Business and Financial Operations Occupations	8,290,731	33,210,983	4.8%	
Computer and Mathematical Occupations	4,283,817	20,067,675	2.9%	
Architecture and Engineering Occupations	2,573,767	12,812,260	1.9%	
Life, Physical, and Social Science Occupations	907,683	4,294,966	0.6%	
Community and Social Service Occupations	1,167,878	7,156,326	1.0%	
Legal Occupations	520,072	4,659,370	0.7%	
Education, Training, and Library Occupations	1,407,548	12,134,228	1.8%	
Arts, Design, Entertainment, Sports, and Media Occupations	2,122,793	9,766,670	1.4%	
Healthcare Practitioners and Technical Occupations	4,132,156	38,777,284	5.7%	
Healthcare Support Occupations	1,540,604	28,300,720	4.1%	
Protective Service Occupations	651,463	7,377,864	1.1%	
Food Preparation and Serving Related Occupations	3,670,975	68,644,547	10.0%	
Building and Grounds Cleaning and Maintenance	2,118,482	21,712,311	3.2%	
Personal Care and Service Occupations	1,373,703	14,691,196	2.1%	
Sales and Related Occupations	8,488,583	84,341,624	12.3%	
Office and Administrative Support Occupations	19,330,700	101,129,136	14.8%	
Farming, Fishing, and Forestry Occupations	163,558	2,534,908	0.4%	
Construction and Extraction Occupations	3,091,109	32,378,260	4.7%	
Installation, Maintenance, and Repair Occupations	4,690,780	28,759,476	4.2%	
Production Occupations	6,987,796	54,476,434	7.9%	
Transportation and Material Moving Occupations	6,192,045	64,208,551	9.4%	

Appendix A: Occupations by Quintile

3-digit SOC code	*SOC Title*	*Average ln(wage)*	*Cummulative percentage of employment*	*Occupation Quintile*
359	Other Food Preparation and Serving Related Workers	1.97	1.1%	1
353	Food and Beverage Serving Workers	1.98	6.7%	1
393	Entertainment Attendants and Related Workers	2.01	7.2%	1
352	Cooks and Food Preparation Workers	2.08	9.6%	1
452	Agricultural Workers	2.08	9.9%	1
412	Retail Sales Workers	2.11	17.4%	1
392	Animal Care and Service Workers	2.11	17.5%	1
372	Building Cleaning and Pest Control Workers	2.12	19.7%	1
311	Nursing, Psychiatric, and Home Health Aides	2.14	22.7%	2
516	Textile Apparel and Furnishings Workers	2.16	23.2%	2
536	Other Transportation Workers	2.17	23.5%	2
399	Other Personal Care and Service Workers	2.19	24.3%	2
396	Baggage Porters Bellhops and Concierges	2.22	24.4%	2
395	Personal Appearance Workers	2.23	24.8%	2
537	Material Moving Workers	2.23	30.3%	2
397	Tour and Travel Guides	2.23	30.3%	2
373	Grounds Maintenance Workers	2.24	31.0%	2
339	Other Protective Service Workers	2.25	32.0%	2
513	Food Processing Workers	2.25	32.6%	2
432	Communications Equipment Operators	2.29	32.8%	2
259	Other Education, Training, and Library Occupations	2.31	33.1%	2
473	Helpers Construction Trades	2.32	33.3%	2
453	Fishing and Hunting Workers	2.34	33.3%	2
517	Woodworkers	2.36	33.6%	2
439	Other Office and Administrative Support Workers	2.36	36.4%	2
512	Assemblers and Fabricators	2.39	38.0%	2
434	Information and Record Clerks	2.40	42.2%	3
519	Other Production Occupations	2.42	44.6%	3
319	Other Healthcare Support Occupations	2.43	45.8%	3
351	Supervisors of Food Preparation and Serving Workers	2.46	46.6%	3
433	Financial Clerks	2.49	49.2%	3
533	Motor Vehicle Operators	2.49	52.2%	3
435	Material, Recording, Scheduling, Dispatching, and Distributing Workers	2.49	53.5%	3
332	Fire Fighting and Prevention Workers	2.49	53.5%	3
515	Printing Workers	2.51	53.7%	3
454	Forest Conservation and Logging Workers	2.53	53.8%	3
253	Other Teachers and Instructors	2.55	54.1%	3
252	Preschool, Primary, Secondary, and Special Education School Teachers	2.55	54.7%	3
514	Metal Workers and Plastic Workers	2.56	56.3%	3

3-digit SOC code	SOC Title	Average ln(wage)	Cummulative percentage of employment	Occupation Quintile
436	Secretaries and Administrative Assistants	2.57	58.9%	3
394	Funeral Service Workers	2.57	58.9%	3
419	Other Sales and Related Workers	2.58	59.6%	3
391	Supervisors of Personal Care and Service Workers	2.58	59.8%	3
333	Law Enforcement Workers	2.59	59.8%	3
211	Counselors, Social Workers, and Other Community and Social Service Specialists	2.60	60.8%	4
371	Supervisors of Building and Grounds Cleaning and Maintenance Workers	2.63	61.0%	4
493	Vehicle and Mobile Equipment Mechanics, Installers, and Repairers	2.63	62.2%	4
499	Other Installation Maintenance and Repair Occupations	2.65	64.3%	4
292	Health Technologists and Technicians	2.67	66.3%	4
212	Religious Workers	2.67	66.4%	4
312	Occupational Therapy and Physical Therapist Assistants and Aides	2.68	66.5%	4
475	Extraction Workers	2.69	66.7%	4
472	Construction Trades Workers	2.71	70.4%	4
274	Media and Communication Equipment Workers	2.71	70.5%	4
474	Other Construction and Related Workers	2.73	70.7%	4
451	Supervisors of Farming, Fishing, and Forestry Workers	2.73	70.7%	4
411	Supervisors of Sales Workers	2.75	72.0%	4
271	Art and Design Workers	2.76	72.4%	4
272	Entertainers and Performers, Sports and Related Workers	2.76	72.8%	4
331	Supervisors of Protective Service Workers	2.77	72.8%	4
194	Life, Physical, and Social Science Technicians	2.77	73.0%	4
299	Other Healthcare Practitioners and Technical Occupations	2.78	73.1%	4
254	Librarians, Curators, and Archivists	2.79	73.1%	4
492	Electrical and Electronic Equipment Mechanics, Installers, and Repairers	2.81	73.6%	4
232	Legal Support Workers	2.83	73.9%	4
531	Supervisors of Transportation and Material Moving Workers	2.87	74.2%	4
431	Supervisors of Office and Administrative Support Workers	2.89	75.3%	4
535	Water Transportation Workers	2.91	75.3%	4
173	Drafters, Engineering Technicians, and Mapping Technicians	2.91	75.9%	4
273	Media and Communication Workers	2.93	76.3%	4
511	Supervisors of Production Workers	2.98	76.8%	4

3-digit SOC code	SOC Title	Average ln(wage)	Cummulative percentage of employment	Occupation Quintile
534	Rail Transportation Workers	2.98	76.9%	4
413	Sales Representatives: Services	2.99	78.4%	4
518	Plant and System Operators	3.02	78.5%	4
414	Sales Representatives: Wholesale and Manufacturing	3.06	80.1%	5
491	Supervisors of Installation, Maintenance, and Repair Workers	3.08	80.4%	5
471	Supervisors of Construction and Extraction Workers	3.10	80.8%	5
131	Business Operations Specialists	3.10	83.8%	5
195	Occupational Health and Safety Specialists and Technicians	3.11	83.8%	5
193	Social Scientists and Related Workers	3.18	83.9%	5
132	Financial Specialists	3.18	85.8%	5
171	Architects, Surveyors, and Cartographers	3.18	86.0%	5
251	Postsecondary Teachers	3.20	86.4%	5
532	Air Transportation Workers	3.23	86.6%	5
151	Computer Specialists	3.30	89.4%	5
192	Physical Scientists	3.32	89.6%	5
191	Life Scientists	3.32	89.7%	5
119	Other Management Occupations	3.34	91.1%	5
152	Mathematical Science Occupations	3.35	91.2%	5
291	Health Diagnosing and Treating Practitioners	3.38	94.8%	5
172	Engineers	3.43	96.0%	5
111	Top Executives	3.61	97.8%	5
113	Operations Specialties Managers	3.63	99.0%	5
112	Advertising, Marketing, Promotions, Public Relations, and Sales Managers	3.67	99.6%	5
231	Lawyers, Judges, and Related Workers	3.75	100.0%	5

Appendix B: Dropping Imputations

General practice in the wage inequality literature based on the Current Population Survey, such as Lemieux (2006), is to drop imputed data in the analysis. However, the imputations in the Occupational Employment and Wage Statistics microdata are an integral part of the estimation strategy for official publications based on these data. These imputations are constructed with a great deal of information on employer location, industry, and size from the Quarterly Census of Employment and Wages, using nearest-neighbor matching with separate procedures for employment and wage variables. The estimation weights assume the inclusion of the imputed data; the imputation procedures are essentially more detailed weights on non-imputed data. However, in this Appendix, I check that the main results in this paper are robust to dropping imputed data.

As shown below, the results are largely consistent with those in tables 2, 3, and 4.

- Table B1, like Table 2, shows that by each measure of occupational homogeneity, overall and for all quintiles of the occupational distribution, with and without additional controls, greater occupational homogeneity is associated with higher wages—except for the top quintile of occupations by the predicted variance of wages between occupations measure, with the additional controls.
- Table B2, like Table 3, shows overall increases in occupational homogeneity by the Herfindahl-Hirschman measure, as well as increases by this measure in the lowest and highest paid quintile of occupations with and without controls, and for every quintile with controls except the 4th quintile. This Table also shows little evidence of increases in occupational homogeneity by the partial predicted variance of wages measure overall or for the lowest-paid four quintiles of the occupational distribution, with and without additional controls. It does show decreases in this measure (increased occupational homogeneity) for the highest-paid quintile of the occupational distribution.
- Table B3, like Table 4, shows that changes in the occupational homogeneity of establishments from 2004 to 2016 contribute substantially to increases in overall ln(wage) variance between these two reference dates.

Appendix Table B1: Regressions of log wages on measures of Occupational Homogeneity, dropping imputed data

Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment
All Occupations (57,637,136 observations)		
With only date fixed effects	-0.542	1.878
	(0.002)	(0.008)
All Controls	-0.073	0.118
	(0.001)	(0.004)
Lowest-paid quintile of occupations (5,031,647 observations)		
With only date fixed effects	-0.208	0.725
	(0.002)	(0.006)
All Controls	-0.097	0.242
	(0.002)	(0.006)
Second quintile of occupations (7,302,459 observations)		
With only date fixed effects	-0.221	0.789
	(0.002)	(0.008)
All Controls	-0.104	0.453
	(0.002)	(0.006)
Middle quintile of occupations (13,674,543 observations)		
With only date fixed effects	-0.149	0.599
	(0.003)	(0.006)
All Controls	-0.094	0.411
	(0.002)	(0.006)
Fourth quintile of occupations (12,869,598 observations)		
With only date fixed effects	-0.197	0.578
	(0.002)	(0.007)
All Controls	-0.060	0.207
	(0.002)	(0.008)
Highest-paid quintile of occupations (18,758,889 observations)		
With only date fixed effects	-0.122	0.122
	(0.005)	(0.013)
All Controls	-0.021	-0.584
	(0.004)	(0.011)

Notes: These are coefficients α from regressions of the form $\ln(wage_{ijt}) = \alpha OccHomogeneity_{jt} + \beta X_{ijt} + \varepsilon_{ijt}$, where X includes reference date fixed effects, occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, state fixed effects, and establishment size (using fixed effects for establishment size classes as well as continuous establishment size). Regressions are at the establishment-occupation-wage interval level, weighted by employment. Normalized Herfindahl indices of Occupational Homogeneity for the establishment are calculated at the minor occupational group level, and are normalized for the number of employees in the establishment. Partial Predicted Variances of Wages for the establishment are based on employment by minor occupational group within the establishment, and do not include predicted within-occupational group wage variation. Clustered standard errors in parentheses.

Appendix Table B2: Change in mean values of Occupational Homogeneity over time, dropping imputed data

Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment
All Occupations (57,637,136 observations)		
Raw Trend	0.0050	0.0022
	(0.0006)	(0.0002)
All Controls	0.0024	0.0014
	(0.0004)	(0.0001)
Lowest-paid quintile of occupations (5,031,647 observations)		
Raw Trend	0.0058	0.0026
	(0.0011)	(0.0003)
All Controls	0.0047	0.0032
	(0.0007)	(0.0002)
Second quintile of occupations (7,302,459 observations)		
Raw Trend	0.0293	-0.0019
	(0.0013)	(0.0003)
All Controls	0.0018	0.0026
	(0.0007)	(0.0002)
Middle quintile of occupations (13,674,543 observations)		
Raw Trend	0.0061	0.0028
	(0.0008)	(0.0002)
All Controls	-0.0013	0.0036
	(0.0006)	(0.0002)
Fourth quintile of occupations (12,869,598 observations)		
Raw Trend	-0.0065	0.0042
	(0.0009)	(0.0002)
All Controls	-0.0001	0.0019
	(0.0006)	(0.0001)
Highest-paid quintile of occupations (18,758,889 observations)		
Raw Trend	0.0022	-0.0030
	(0.0010)	(0.0004)
All Controls	0.0041	-0.0036
	(0.0008)	(0.0002)

Note: These are coefficients α from regressions of the form $Occupation\ Homogeneity_{ijt} = \alpha Decades_t + \beta X_{ijt} + \varepsilon_{ijt}$, where Decades measures time in decades since 2004 and X_{ijt} includes occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, establishment size class, continuous establishment size, and state of location. Regressions are at the establishment-occupation-wage interval level, weighted by employment. Occupations (at the minor occupational group level) found in each quintile are listed in Appendix A. Normalized Herfindahl indices of Occupational Homogeneity for the establishment are calculated at the minor occupational group level, and are normalized for the number of employees in the establishment. Partial Predicted Variances of Wages for the establishment are based on employment by minor occupational group within the establishment, and do not include predicted within-occupational group wage variation. Clustered standard errors are in parentheses.

Appendix Table B3: Decomposition of Changes in real ln(wage) variance from 2004 to 2016, dropping imputed data

			Observations in late period (2016): 8,365,716
			Observations in early period (2004): 8,476,264
Real log wage variance	Coeff.	Percent	Standard Deviations
Overall Variance			
Late period (2016 reference date)	0.3767		0.0006
Counterfactual variance	0.3539		0.0005
Early period (2004 reference date)	0.3588		0.0007
Total change	0.0178	100%	0.0009
of which explained (by compositional change)	0.0227	128%	0.0001
of which unexplained (wage structure change)	-0.0049	-28%	0.0009
Explained (compositional effect)			
Total	0.0227	100%	0.0001
Pure explained	0.0226	99%	0.0002
Specification error	0.0002	1%	0.0000
Components of the pure explained effect			
Industry sector (2-digit NAICS)	0.0025	11%	0.0001
Geography (Census Division)	0.0007	3%	0.0000
Establishment size	-0.0003	-1%	0.0000
Occupation quintiles (defined in Appendix A)	0.0167	74%	0.0003
Normalized Herfindahl measure of establishments	0.0023	10%	0.0002
Partial predicted variance of establishment ln(wages)	0.0006	3%	0.0001
Unexplained (wage structure changes)			
Total	-0.0049	100%	0.0009
Reweighting error	-0.0001	2%	0.0004
Pure unexplained	-0.0048	98%	0.0008

Notes: These are the results of Rios-Avila's implementation of Fortin, Firpo, and Lemiueux's Recentered Influence Function Decomposition. OEWS data with a 2004 reference date was collected from 2001 to 2004; date with a 2016 reference date was collected from 2014 to 2016. Industry here is grouped into 19 supersectors and geography into 7 Census divisions. Establishment size is measured in 9 categories (1-4 employees, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, and 1000+). Quintiles of occupations are defined in Appendix A. Establishment-level normalized Herfindahl-Hirschman indices of occupations are measured with quartiles of the distribution, interacted with occupational quintiles. The predicted variance of ln(wage) for establishments is also divided into quartiles and interacted with occupational quintiles, with an additional dummy variable for low-wage occupations in establishments of less than 100 workers that are in the bottom half of the predicted variance distribution. Standard deviations have not been bootstrapped.

Appendix C: Measuring employer size and homogeneity using EINs instead of establishments

Song, Price, Guvenen, Bloom, and von Wachter (2019) argue that the unit of importance for wage inequality should be the firm and not the establishment. In thinking about occupational homogeneity, some of the reasons for employers to outsource work to other establishments are also reasons to outsource work to other employers entirely. It may be more efficient for even multi-establishment employers to specialize in particular areas of work. Regulatory incentives for multi-establishment employers to specialize in employing workers in a particular part of the wage distribution are less clear. ERISA laws define employers as "controlled groups of corporations" and "entities under common control" in requiring common levels of pension and welfare benefits among most employees in exchange for favorable tax treatment (Perun 2010), and the Affordable Care Act of 2010 extended these provisions by requiring common levels of health care benefits among most employees of businesses with a common owner. However, as Perun notes, "Employers often invent new organizational structures and worker classifications designed to limit participation to favored employees… Regulatory authorities in turn develop complicated rules and regulations designed to prevent this."

This paper focuses on measures of occupational homogeneity at the establishment level because establishments are the sampling units of the OEWS, and the OEWS sampling design often includes some but not all establishments of multi-establishment companies, particularly when there are establishments in geographic areas with fewer establishments available to sample. However, the OEWS microdata can be linked with the EIN (tax-ID) numbers that these establishments submit to the unemployment insurance system, as compiled by the Quarterly Census of Employment and Wages. As discussed extensively in Handwerker and Mason (2013), very large firms may use multiple EINs in the unemployment insurance system, and linking together all of the establishments in these data for very large firms involves a tremendous amount of manual review. Thus, while it is straightforward to recalculate measures of occupational homogeneity at the EIN level and repeat the analyses above, the reader should be cautioned that such EIN-level measures are not true firm-level measures.

OEWS data show that workers in the bottom quintile of occupations were paid more in huge firms than in smaller firms during earlier waves of data collection, but this difference disappeared around November 2013. This is consistent with the finding of Song et. al. that workers with low values of worker fixed effects in very large firms have seen declining wages over time. It is not exactly comparable to Song et. al. because those authors use repeated observations of workers over time to estimate worker fixed effects, an estimation not possible with the OEWS data. However, there is likely a great deal of overlap between workers in typically-low-wage occupations and workers with low fixed effects.

As shown below, the results are largely consistent with those in tables 2 and 3.

- Table C1, like Table 2, shows that by each measure of occupational homogeneity, overall and for all quintiles of the occupational distribution, with and without additional controls, greater occupational homogeneity is associated with higher wages—except for the top quintile of occupations by the predicted variance of wages between occupations measure, with the additional controls.

- Table C2, like Table 4, shows mixed evidence on trends in occupational homogeneity, with more consistent overall evidence of increased homogeneity over time by the Herfindahl-Hirschman measure than by the partial predicted variance of wages measure. Both measures show increased occupational homogeneity over time for the highest-paid quintile of occupations.
- Table C3, does not show the same role for increased occupational homogeneity in increasing overall wage variance between the 2004 and 2016 reference dates for the partial predicted variance of wages measure, but does for the Herfindahl-Hirschman measure of occupational homogeneity.

Appendix Table C1: Regressions of log wages on EIN-level measures of Occupational Homogeneity

Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the EIN	Partial Predicted Variance of Wages for the EIN
All Occupations (94,626,298 observations)		
With only date fixed effects	-0.569	2.060
	(0.011)	(0.027)
All Controls	-0.051	0.123
	(0.002)	(0.007)
Lowest-paid quintile of occupations (7,725,178 observations)		
With only date fixed effects	-0.174	0.677
	(0.006)	(0.015)
All Controls	-0.071	0.189
	(0.005)	(0.010)
Second quintile of occupations (11,782,345 observations)		
With only date fixed effects	-0.230	0.758
	(0.007)	(0.022)
All Controls	-0.083	0.360
	(0.004)	(0.011)
Middle quintile of occupations (22,319,551 observations)		
With only date fixed effects	-0.136	0.538
	(0.007)	(0.027)
All Controls	-0.074	0.338
	(0.003)	(0.009)
Fourth quintile of occupations (20,693,840 observations)		
With only date fixed effects	-0.223	0.676
	(0.007)	(0.022)
All Controls	-0.033	0.167
	(0.003)	(0.012)
Highest-paid quintile of occupations (32,105,384 observations)		
With only date fixed effects	-0.130	0.327
	(0.009)	(0.026)
All Controls	-0.003	-0.361
	(0.005)	(0.014)

Notes: These are coefficients α from regressions of the form $\ln(wage_{ijt}) = \alpha OccHomogeneity_{jt} + \beta X_{ijt} + \varepsilon_{ijt}$, where Date X includes reference date fixed effects, occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, state fixed effects, and EIN size (using fixed effects for EIN size classes as well as continuous EIN size). Regressions are at the establishment-occupation-wage interval level, weighted by employment. Normalized Herfindahl indices of Occupational Homogeneity for the EIN are calculated at the minor occupational group level, and are normalized for the number of employees in the EIN. Partial Predicted Variances of Wages for the EIN are based on employment by minor occupational group within the EIN, and do not include predicted within-occupational group wage variation. Clustered standard errors in parentheses.

Appendix Table C2: Change in mean values of EIN-level Occupational Homogeneity over time

Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the EIN	Partial Predicted Variance of Wages for the EIN
All Occupations (94,626,298 observations)		
Raw Trend	0.0052	0.0019
	(0.0011)	(0.0004)
All Controls	0.0001	0.0024
	(0.0008)	(0.0003)
Lowest-paid quintile of occupations (7,725,178 observations)		
Raw Trend	0.0047	0.0009
	(0.0021)	(0.0006)
All Controls	-0.0019	0.0034
	(0.0019)	(0.0006)
Second quintile of occupations (11,782,345 observations)		
Raw Trend	0.0280	-0.0009
	(0.0022)	(0.0006)
All Controls	0.0008	0.0037
	(0.0014)	(0.0004)
Middle quintile of occupations (22,319,551 observations)		
Raw Trend	0.0067	0.0025
	(0.0011)	(0.0003)
All Controls	-0.0006	0.0040
	(0.0008)	(0.0002)
Fourth quintile of occupations (20,693,840 observations)		
Raw Trend	-0.0083	0.0047
	(0.0011)	(0.0003)
All Controls	-0.0019	0.0025
	(0.0007)	(0.0002)
Highest-paid quintile of occupations (32,105,384 observations)		
Raw Trend	0.0010	-0.0018
	(0.0010)	(0.0005)
All Controls	0.0019	-0.0014
	(0.0008)	(0.0004)

Note: These are coefficients α from regressions of the form $Occupation\ Homogeneity_{ijt} = \alpha Decades_t + \beta X_{ijt} + \varepsilon_{ijt}$, where Decades measures time in decades since 2004 and X_{ijt} includes occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, EIN size class, continuous EIN size, and state of location. Regressions are at the establishment-occupation-wage interval level, weighted by employment. Occupations (at the minor occupational group level) found in each quintile are listed in Appendix A. Normalized Herfindahl indices of Occupational Homogeneity for the establishment are calculated at the minor occupational group level, and are normalized for the number of employees in the EIN. Partial Predicted Variances of Wages for the establishment are based on employment by minor occupational group within the EIN, and do not include predicted within-occupational group wage variation. Clustered standard errors are in parentheses.

Appendix Table C3: Decomposition of Changes in real ln(wage) variance from 2004 to 2016, using EIN-level measures of Occupational Homogeneity and employer size

	Observations in late period (2016): 14,015,811		
	Observations in early period (2004): 13,012,511		
Real log wage variance	Coeff.	Percent	Standard Deviations
Overall Variance			
Late period (2016 reference date)	0.3803		0.0005
Counterfactual variance	0.3594		0.0004
Early period (2004 reference date)	0.3621		0.0006
Total change	0.0181	100%	0.0007
of which explained (by compositional change)	0.0209	115%	0.0001
of which unexplained (wage structure change)	-0.0027	-15%	0.0007
Explained (compositional effect)			
Total	0.0209	100%	0.0001
Pure explained	0.0207	99%	0.0001
Specification error	0.0002	1%	0.0000
Components of the pure explained effect			
Industry sector (2-digit NAICS)	0.0019	9%	0.0001
Geography (Census Division)	0.0004	2%	0.0000
EIN size	-0.0002	-1%	0.0000
Occupation quintiles (defined in Appendix A)	0.0179	86%	0.0002
Normalized Herfindahl measure of establishments	0.0015	7%	0.0001
Partial predicted variance of establishment ln(wages)	-0.0008	-4%	0.0001
Unexplained (wage structure changes)			
Total	-0.0027	100%	0.0007
Reweighting error	-0.0000	1%	0.0003
Pure unexplained	-0.0027	99%	0.0006

Notes: These are the results of Rios-Avila's implementation of Fortin, Firpo, and Lemiueux's Recentered Influence Function Decomposition. OEWS data with a 2004 reference date was collected from 2001 to 2004; date with a 2016 reference date was collected from 2014 to 2016. Industry here is grouped into 19 supersectors and geography into 7 Census divisions. EIN size is measured in 10 categories (1-4 employees, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1,000-9,999, and 10,000+). Quintiles of occupations are defined in Appendix A. EIN-level normalized Herfindahl-Hirschman indices of occupations are measured with quartiles of the distribution, interacted with occupational quintiles. The predicted variance of ln(wage) for EINs is also divided into quartiles and interacted with occupational quintiles, with an additional dummy variable for low-wage occupations in EINsof less than 100 workers that are in the bottom half of the predicted variance distribution. Standard deviations have not been bootstrapped.

Appendix D: Heterogeneity by state-level unionization rates

One factor which may impact both wages and the organization of production (including the variety of occupations at a workplace) is unionization. The OEWS does not collect information on unionization patterns by employer, but these data include the location of each establishment, and unionization rates vary strongly by state. Thus, I use state-level unionization rates to group observations by whether they were collected in highly unionized states (17-26% of employed workers unionized), middle, and low unionized states (3-9.3% unionized), based on published tables from the Current Population Survey.

Overall, the highest levels of occupational homogeneity are in states with low levels of unionization. This is also true for workers in the lowest-paid quintile of occupations. However, although the lowest levels of occupational homogeneity overall are in states with the highest levels of unionization, for workers in the lowest-paid quintile of occupations, the states with the highest levels of occupational homogeneity have middle levels of unionization.

Following equation (3), the relationships between occupational homogeneity and wages are estimated separately for the most and least unionized groups of states, and these are shown in Table D1. There is no clear pattern of differences between the results by state unionization levels. Whether occupational homogeneity matters more for wages in more or less unionized states varies by the measure of occupational homogeneity, which workers are examined, and whether or not controls are included for establishment characteristics and occupation.

Differences in occupational homogeneity trends between less and more unionized states (following equation (5)) show that establishments are growing more occupationally homogeneous over time in the less-unionized states, relative to the highly unionized states, by the predicted wage variance between-occupations. However, as shown in Table D2, there is no clear pattern of differences between more and less unionized states of trends in the normalized Herfindahl-Hirschman measure of occupational homogeneity

Table D1: Regressions of log wages on measures of Occupational Homogeneity, by Unionization group

	Most Unionized States		Least Unionized States	
Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment
All Occupations	**(32,864,292 observations)**		**(31,742,453 observations)**	
With only date fixed effects	-0.545	1.875	-0.517	1.800
	(0.003)	(0.012)	(0.003)	(0.009)
All Controls	-0.055	0.085	-0.065	0.093
	(0.002)	(0.006)	(0.002)	(0.005)
Lowest-paid quintile of occupations	**(2,610,856 observations)**		**(2,654,471 observations)**	
With only date fixed effects	-0.176	0.622	-0.168	0.594
	(0.003)	(0.009)	(0.002)	(0.008)
All Controls	-0.062	0.149	-0.079	0.172
	(0.003)	(0.010)	(0.002)	(0.008)
Second quintile of occupations	**(4,165,829 observations)**		**(4,003,654 observations)**	
With only date fixed effects	-0.223	0.768	-0.219	0.658
	(0.003)	(0.011)	(0.003)	(0.010)
All Controls	-0.082	0.346	-0.081	0.363
	(0.003)	(0.008)	(0.003)	(0.008)
Middle quintile of occupations	**(7,813,498 observations)**		**(7,508,503 observations)**	
With only date fixed effects	-0.144	0.517	-0.096	0.423
	(0.004)	(0.009)	(0.004)	(0.009)
All Controls	-0.077	0.316	-0.075	0.320
	(0.003)	(0.008)	(0.003)	(0.008)
Fourth quintile of occupations	**(6,983,355 observations)**		**(7,181,134 observations)**	
With only date fixed effects	-0.149	0.336	-0.247	0.726
	(0.004)	(0.011)	(0.003)	(0.009)
All Controls	-0.037	0.144	-0.072	0.218
	(0.003)	(0.011)	(0.003)	(0.009)
Highest-paid quintile of occupations	**(11,290,754 observations)**		**(10,394,691 observations)**	
With only date fixed effects	-0.115	0.184	-0.103	0.079
	(0.008)	(0.021)	(0.006)	(0.012)
All Controls	-0.009	-0.439	-0.025	-0.455
	(0.005)	(0.014)	(0.005)	(0.013)

Notes: These are coefficients α from regressions of the form $\ln(wage_{ijt}) = \alpha OccHomogeneity_{jt} + \beta X_{ijt} + \varepsilon_{ijt}$, where X includes reference date fixed effects, occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, state fixed effects, and establishment size (using fixed effects for establishment size classes as well as continuous establishment size). Regressions are at the establishment-occupation-wage interval level, weighted by employment. Clustered standard errors in parentheses.

Table D2: Change in mean Occupational Homogeneity, 2004-2019, by Unionization

	Most Unionized States		Least Unionized States	
Occupational Homogeneity Variable	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment	Normalized Herfindahl of Occupational Homogeneity for the establishment	Partial Predicted Variance of Wages for the establishment
All Occupations	**(32,864,292 observations)**		**(31,742,453 observations)**	
Raw Trend	0.0117	0.0015	0.0112	-0.0012
	(0.0009)	(0.0003)	(0.0008)	(0.0003)
All Controls	0.0036	0.0022	0.0056	-0.0009
	(0.0006)	(0.0002)	(0.0005)	(0.0002)
Lowest-paid quintile of occupations	**(2,610,856 observations)**		**(2,654,471 observations)**	
Raw Trend	0.0122	0.0015	0.0177	-0.0022
	(0.0017)	(0.0004)	(0.0016)	(0.0004)
All Controls	0.0035	0.0039	0.0111	0.0000
	(0.0011)	(0.0003)	(0.0010)	(0.0003)
Second quintile of occupations	**(4,165,829 observations)**		**(4,003,654 observations)**	
Raw Trend	0.0416	-0.0027	0.0259	-0.0023
	(0.0021)	(0.0005)	(0.0017)	(0.0004)
All Controls	0.0062	0.0027	0.0047	0.0007
	(0.0011)	(0.0003)	(0.0011)	(0.0002)
Middle quintile of occupations	**(7,813,498 observations)**		**(7,508,503 observations)**	
Raw Trend	0.0074	0.0031	0.0133	-0.0006
	(0.0011)	(0.0003)	(0.0013)	(0.0003)
All Controls	-0.0014	0.0049	0.0021	0.0008
	(0.0008)	(0.0002)	(0.0009)	(0.0002)
Fourth quintile of occupations	**(6,983,355 observations)**		**(7,181,134 observations)**	
Raw Trend	-0.0029	0.0043	0.0013	0.0008
	(0.0014)	(0.0004)	(0.0011)	(0.0003)
All Controls	0.0001	0.0032	0.0029	-0.0005
	(0.0009)	(0.0002)	(0.0008)	(0.0002)
Highest-paid quintile of occupations	**(11,290,754 observations)**		**(10,394,691 observations)**	
Raw Trend	0.0056	-0.0030	0.0039	-0.0056
	(0.0016)	(0.0007)	(0.0012)	(0.0005)
All Controls	0.0053	-0.0025	0.0038	-0.0049
	(0.0010)	(0.0003)	(0.0010)	(0.0003)

Note: These are coefficients α from regressions of the form $Occupation\ Homogeneity_{ijt} = \alpha Decades_t + \beta X_{ijt} + \varepsilon_{ijt}$, where Decades measures time in decades since 2004 and X_{ijt} includes occupation fixed effects at the 6-digit SOC level, industry fixed effects at the 4-digit NAICS level, establishment size class, continuous establishment size, and state of location. Regressions are at the establishment-occupation-wage interval level, weighted by employment. Standard errors are in parentheses.