Increased Concentration of Occupations, Outsourcing, and Growing Wage Inequality in the United States

Elizabeth Weber Handwerker
U.S. Bureau of Labor Statistics

July, 2018

Abstract: This paper develops measures of the concentration of occupational employment within employers to study the extent of outsourcing in the economy as a whole. Findings are threefold. First, wages are strongly related to the occupational concentration of workers within establishments. Workers in establishments more concentrated in occupations are paid lower wages. This relationship holds even after controlling for workers' own occupations and observable employer characteristics, and has been increasing somewhat during 2002-2016. Second, during this same period, after controlling for observable establishment characteristics, workers in low-wage occupations saw their employing establishments became more concentrated in the mix of occupations employed. Third, occupational distributions can explain a substantial amount of variation in wage levels between employers, and changes in the occupations and occupational concentration of workers can more than explain the observed increase in overall and between-establishment wage inequality during this time period. Measures of occupational concentration that incorporate the presence or absence of workers in high-wage occupations from an establishment appear to be particularly important in explaining the growth of wage inequality. Thus, the growing separation of workers in low-wage occupations into different employers from workers in high-wage occupations appears to be 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 of Equitable Growth, and the Bureau of Labor Statistics for comments, with special thanks to Matt Dey, Katherine Abraham, Erica Groshen, Sue Houseman, David Levine, Anne Polivka, Alex Bryson, David Piccone, 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 to the development of this paper. Research assistance was provided by Lowell Mason. 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.

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 clerical work, and the increased potential for imports to replace domestic labor. This paper examines an additional source of growing wage inequality: the changing distribution of occupations between establishments 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 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 that the increasing variability is 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, 2016),[1] and that 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, 2016). The results in this paper show that occupational specialization—a specific form of worker sorting—is a key explanation for the growth in between employer wage inequality: a growing trend of workers in low-wage occupations employed at different employers from other 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 differing portions of the worker skill distribution[4]. As wages for different kinds of work become less equal, employers face 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.

---

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

[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.

There are other potential reasons for businesses to outsource work, such as increasing the ability to smooth workload and economies of scale available to providers of specialized services (Abraham and Taylor, 1996). However, to the extent that labor cost savings and avoiding paying efficiency wages or rents when market wages are low for particular types of low-skill work drive outsourcing decisions, this may lead to establishments specializing in high and low-wage work, and result in growing wage inequality across establishments. Recent work by Goldschmidt and Schmeider (2015) shows labor cost savings to be a primary driver of outsourcing in Germany, as outsourced workers lose firm-specific rents. In three well-defined occupational categories, they find that losses of these 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 had to focus on particular industries or occupations associated with performing support tasks for 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, using the concentration of occupations at an employer, as measured in the detailed microdata of the Occupational Employment 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 are outsourcing work to avoid monitoring, hiring, or other costs for occupations in which they have less expertise, there will be less variety in the number of occupations they employ. However, when businesses are outsourcing 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 these changes in outsourcing and occupational concentration are compared with the effect 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 concentration of workers within establishments. Workers in establishments that are more concentrated in occupations overall 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 generally low-wage occupations. Second, from November 2002 through November 2016, occupational concentration has slightly increased for workers in typically low-wage occupations, after controlling for other employer characteristics. Third, changes in the distribution of occupational concentration are related to the growth in private-sector wage inequality observed in the data over the 2002/2003 - 2016 time period. Using these occupational concentration measures, it is possible to explain as much as 144% of overall wage inequality growth among workers and 99% of wage inequality growth between their employing establishments. Changes in the distributions of occupations of workers and the industries, establishment sizes, and the geography of their employing establishments can explain no more than 124% of overall wage inequality growth and no more than 73% of wage inequality growth between establishments over this period. Among the measures of occupational concentration, measures based on the distribution of occupations by their wage levels matter more for wage inequality and its growth than measures that ignore wage differences between occupations, and the presence of high-wage occupations in an establishment is particularly important.

The paper is organized as follows: Section II describes the main measures of occupational concentration. Section III describes relationships between occupational concentration and wages. Section IV describes trends in measured occupational concentration. Section V describes the impact of occupation and occupational concentration measures on wage inequality between establishments and wage inequality over time. Section VI repeats much of the main analysis for larger measures of employers than the establishments. Section VII concludes.

II. Measuring Occupational Concentration

Occupational concentration is the variety of occupations employed at a place of business. It is a descriptive measure of the variety of tasks performed by the employer, separate from describing the tasks performed by individual employees (their occupations), the type of work done at the establishment (the industry) or its size. 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 concentration is intended as a measure of the variety of the type of work done in establishments throughout the economy. In this section, measures of occupational concentration are defined, and evidence is presented to show that occupational concentration matters for wages and displays a changing distribution over time.

Two measures of occupational concentration within establishments are constructed: a more general occupational concentration measure across all occupations, regardless of whether they are high or low paying occupations, and a measure that explicitly models the variation in wages due to the distribution of occupations for each establishment. We define the general occupational concentration across all occupations in an establishment with a Herfindahl index:

$$H_{jt} = \sum_{k=1}^{98} \left(\frac{Employment\ in\ Detailed\ Occupation_{kjt}}{Total\ Employment_{jt}} \right)^2$$

This index uses the 100 minor occupational categories at the 3-digit level of the Standard Occupational Classification system.[6] It varies from 1/100 (equal representation of all

[6] Handwerker and Spletzer (2015) studied this type of general occupational concentration with Herfindahl 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.

occupations) to 1 (perfect concentration). Increased occupational concentration at the establishment level, as measured in this index, indicates that employers are becoming more specialized and are outsourcing work to other employers. Examining trends in the Herfindahl index of occupational concentration should give an indication of whether occupations are becoming more concentrated at the establishments throughout the U.S. economy. However, this index cannot distinguish between outsourcing work that is usually paid very similar wages to other occupations, and outsourcing work that is usually paid much higher or lower wages than that done by other occupations of the employer.

In contrast, a second index gives a measure of occupational concentration based on the similarity or dissimilarity of typical wages for the occupations each establishment employs. This index is calculated using part of the predicted variance of wages—predicted from the the distribution of employment by occupation for the establishment. Using the average wage for each occupation (for comparability with the index above, we use the 100 minor occupational categories), we model the wage paid by employer j to worker i in occupation o as $\widehat{w_{ij}} = \overline{w_o} + \varepsilon_{ij}$, where $\overline{w_o}$ is the mean wage for employees in occupation o and ε_{ij} is distributed log-normally, with mean 0 and standard deviation σ_o. Similarly, we model the conditional estimated variance of wages for employer j (conditional on the occupational distribution of the workers) as $\widehat{V_j} = \frac{\sum_i \left(\widehat{w_{ij}} - \widehat{\widehat{w_j}}\right)^2}{n_j}$. Using only the occupations of the workers in employer j, the mean wage for employer j is $\widehat{\widehat{w_j}} = \frac{\sum_{all\ occupations\ o} \sum_{i \in o} \overline{w_o}}{n_j}$. Then, $\widehat{V_j} = \frac{\sum_{all\ occupations} \sum_{i \in occupation} \left(\widehat{w_{ij}}^2 - 2\widehat{w_{ij}}\widehat{\widehat{w_j}} + \widehat{\widehat{w_j}}^2\right)}{n_j}$

$$= \frac{\sum_o \sum_{i \in o} \left(\left(\overline{w_o} - \widehat{\widehat{w_j}}\right)^2 + \varepsilon_{ij}^2\right)}{n_j} = \frac{\sum_{all\ occupations} \left[\sum_{i \in o} \left(\overline{w_o} - \widehat{\widehat{w_j}}\right)^2 + \sum_{i \in o} \varepsilon_{ij}^2\right]}{n_j} = \frac{\sum_{all\ occupations} n_o \left[\left(\overline{w_o} - \widehat{\widehat{w_j}}\right)^2 + \sigma_o^2\right]}{n_j}$$

$$= \frac{\sum_{all\ occupations} n_o \left[\left(\overline{w_o} - \widehat{\widehat{w_j}}\right)^2\right]}{n_j} + \frac{\sum_{all\ occupations} n_o \sigma_o^2}{n_j}.$$

This is the variance of log wages for employer j predicted from the composition of occupations employed at employer j, and it is the sum of two parts: the sum of the variation in the means of wages for these occupations and the average of the within-occupation log wage variance for these occupations. This overall predicted variance of wages is our second index of occupational concentration (although the first, "between-occupations" portion of this index is often examined as a third index).

This paper uses the microdata of the Occupational Employment Statistics (OES) Survey for the private sector in the United States for 2002 to 2016,[7] reweighted to match the detailed industry and employer size distribution of the Quarterly Census of Employment and Wages for the appropriate quarter. The OES survey is designed to measure occupational employment and wages in the United States by geography and industry, covering all establishments in the United States except for those in agriculture, private households, and unincorporated self-employed workers without employees. For hundreds of thousands of establishments per year, these microdata record the number of employees for each wage interval within detailed occupation

[7] An earlier version of this paper used these microdata for 1999 to 2015. However, as described by Abraham and Spletzer (2010), many first-line supervision occupations in establishments of less than 500 workers were erroneously coded as managerial occupations during 1999-2001. Thus, data for these earlier years have been omitted from all estimates.

categories. More details about the survey and the reweighting procedure used can be found in the Data Appendix.[8]

Summary statistics are found in Table 1. The average worker has an inflation-adjusted was of \$16.20/hour, or a ln(wage) of 2.56, and is observed in an establishment with a measured ln(wage) variance of 0.154. The average worker's establishment has Herfindahl index of occupational concentration of 0.408, and a predicted variance of ln(wages) estimated from its occupational composition of 0.270, of which 0.106 is due to the between-occupations component of the predicted variance. It is unsurprising that the predicted ln(wage) variance based only on the occupations employed at the establishment is higher than the measured ln(wage) variance, both because wages vary geographically (and the predicted variances do not take geographic variation into account), and because of the large literature describing the impact of employers on wages. This table also presents summary information on the composition of occupations and industries in the reweighted data.

Table 2 compares these three measures of establishment-level occupational concentration for several occupation-industry groups studied as examples of outsourcing by Abraham and Taylor (1996); Dube and Kaplan (2010); Dey, Houseman, and Polivka (2010); and Goldschmidt and Schmeider (2015): food preparation and serving, janitors, security guards, truck drivers, accountants, computer occupations, engineers, 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 that they provide these services to. Table 2 shows that for every single one of these example occupations, Herfindahl indices for the employers of these workers are higher, on average, indicating greater occupational concentration, when they are employed in their specialty industry than when they are employed in other industries. Moreover, for every occupation except lawyers (the smallest, highest paid, and most concentrated in its specialty industry), the predicted variances of wages based on the occupational distribution of their employers—and the between occupations components of these predicted variances—are lower, on average, indicating greater occupational concentration, when they are employed in their specialty industry than when they are employed in other industries. Thus, the three measures of occupational concentration measures defined in this section, designed to measure outsourcing across all occupations and industries, indicate greater occupational concentration in the relevant industries to which workers are outsourced, for specific occupations studied in the outsourcing case-study literature.

III: Relationships between Measured Occupational Concentration and Wages

Measures of occupational concentration are strongly and significantly related to wages—especially for workers in typically low-wage occupations. For the example occupations of Table 2, food preparation and service, janitors, security guards, engineers, and lawyers earn considerably lower wages, on average, in the outsourced specialty industries described above

[8] $\overline{w_o}$ and ${\sigma_o}^2$ are estimated separately for each for each occupation in each time period to calculate the variance of log wages predicted for each employer based on the occupations employed in that time period.

than in other industries. Meanwhile, truck drivers, accountants, and computer occupation earn higher wages, on average, in specialty industries than in other industries.

Looking at all workers in all industries, the relationships between occupational concentration and wages are shown graphically in Figure 1, which plots the average value of ln(wage) by occupational concentration. In this figure, levels of occupational concentration are rounded to the nearest hundredth and each occupation-wage-interval-establishment observation is assigned to a category for that interval. The average impact of each level of concentration on log wages is given by the coefficient α_m in Equation 1 below, for observations in occupational wage interval *i* at establishment *j* and time *t*.

$$\text{(1)} \qquad \ln\left(wage_{ijt}\right) = \sum \alpha_m I\left(m - 1 < Occ\ Concen_{jt} \leq m\right) + \delta X_{ijt} + \varepsilon,$$

Observations with negative (positive) values of occupational concentration are less (more) occupationally concentrated than the U.S. as a whole. This figure shows both the unadjusted coefficients (with no controls for establishment characteristics X), and the adjusted regression coefficients for each level after controlling for survey date fixed effects, occupation fixed effects, industry fixed effects, state fixed effects, and establishment size.[9]

This figure clearly shows that increasing Herfindahl indices of occupational concentration are associated with lower wages, while increasing values of the predicted variance of wages based on the occupational fractions of high wage workers in an establishment are associated with higher wages. These relationships remain (although they are greatly weakened) after controlling for observable characteristics.

The relationships between wages across all occupations and continuous measures of occupational concentration are shown in Table 3, using regressions of the form

$$\text{(2)} \qquad Ln(wage_{ijt}) = \alpha OccConcen_{jt} + \beta OccConcen_{jt} {*} Date + \delta X_{ijt} + \varepsilon$$

The first rows of Table 3 give estimates of the impact of occupational concentration with no additional controls X. These estimates clearly show that increased occupational concentration is associated with lower wages. Estimates of the coefficients β, estimated in decades of time since November 2002, indicate that all these relationships have significantly strengthened over time. Lower sections of Table 3 give estimates with X variables added. These detailed controls reduce the magnitude of the relationship between occupational concentration and wages, but all of the estimates of the main effects α maintain the same sign and remain very significant.

Figure 2 repeats the analysis of Figure 1, but separates observations by occupation (at the 3-digit SOC level) into quintiles by their average wages. Appendix A lists the occupations in each quintile by average ln(wage), while Table 1 tabulates the employment in each quintile and the number of establishments containing these occupations. There are relatively few, large,

[9] Establishment size effects are modeled with both fixed effects for establishment size classes and continuous establishment size.

occupations in the lowest-paid quintile, and a greater number of small occupations in the highest-paid quintile, with roughly a fifth of all (weighted) employment in each.[10]

Figure 2 shows that the negative relationship between occupational concentration and wages is present in occupations of every wage level, but for the quintile of occupations generally paid the highest wages, this relationship reverses once controls are added for industry, establishment size, state, and own-occupation. By contrast, for the occupations in the lowest-paid four quintiles, the relationship between occupational concentration and wages is not fully explained by these other variables.

Table 4 repeats the analysis of Table 3 by occupational groups, using the same quintiles of occupation used in Figure 2. For simplicity of presentation, the coefficients β, on occupational-concentration*time are not shown here, although they are included in the regressions. The overall relationship between occupational concentration and wages (particularly for the predicted-variance measures of occupational concentration, and after controlling for other observed characteristics) is clearly driven by workers in typically low-wage occupations. For workers in low-wage occupations, the estimates α of the relationships between wages and all measures of occupational concentration are particularly strong, showing a negative relationship between occupational concentration and wages. However, estimates of β indicate that the impact of occupational concentration on wages in the lowest-wage jobs has been decreasing over time.

However, the relationships β for the change in the impact of occupational concentration over time reverse (at the fourth quintile of occupations for the Herfindahl index; at the second quintile of occupations for the predicted variance measure and its first component), indicating the impact of occupational concentration on wages has been growing over time for workers in these typically middle-wage occupations.

For workers in the topmost quintile of occupation, higher levels of occupational concentration are associated with *lower* wages by the predicted variance of wages measures, once industry, own-occupation, geographic and size differences are taken into account. Coefficients β for workers in this quintile of occupations show that this relationship has been weakening over time for the Herfindahl measure of concentration, and strengthening over time for the predicted-variance measures of concentration.

Further heterogeneity between occupational concentration and wages—by state-level unionization rates, establishment size, establishment age, industrial sector, and Employer Tax Identification Number (EIN) size, is described in Appendix B.

Together, these results show very strong relationships between occupational concentration and wages. These relationships are only partially explained by occupation and employer characteristics. Moreover, occupational concentration is a particularly important determinant of wages for low-wage workers.

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

IV: Trends in Occupational Concentration Measures

Understanding trends in occupational concentration measures is made more complex by changes in the relative sizes of different occupations during this period. As described by Autor, Katz, and Kearney (2006, 2008), among others, there have been increases in employment typically low-wage and typically high-wage occupations, while employment in many typically middle wage-occupations has fallen. For the occupational quintiles used in Figure 2 and Table 4, employment over time is shown in Figure 3. This figure shows that the same patterns of employment polarization is present in these data: the percentage of employment in the top and bottom quintiles has increased, while the percentage of employment in the middle three quintiles has decreased. This polarization means that ignoring the grouping of employment into establishments entirely, the portion of the variance of ln(wages) for all workers due to variation between occupations is generally increasing—from values of .20 and .21 in the early years of the microdata to .22 and .23 in later years. There is no similar such trend in a version of the Herfindahl index that pools workers across all employers; it varies only between .027 and .028, with no clear time pattern.

Figure 4 plots average values of each employer-level occupational concentration measure for all observations at each OES survey date, with and without adjusting for occupation, detailed industry, size class, establishment size, and state, as the coefficients γ_t from regressions of the form

(3) $$Occupational\ Concen_{jt} = \gamma_t Survey\ Date_t + \delta X_{ijt} + \varepsilon$$

Both with and without these controls, occupational concentration as measured with the Herfindahl index is increasing slightly. As measured by the predicted variance of ln(wages), occupational concentration is concentrating somewhat from November 2002 through November 2009, diluting from November 2009 through November 2015, and concentrating again afterwards. By the between-occupations component of this predicted variance there is a trend towards decreased concentration from November 2002 through May 2012, and then a sharper trend towards increased concentration afterwards. These periods of increased concentration (falling predicted wage variances) at the establishment level occur against the backdrop of the overall polarization of employment, which mechanically leads to increases in the variance of wages between occupations overall.

Figure 5 shows these trends for each quintile of occupations. The pattern for workers in the lowest-paid quintile of occupations is clear: by every measure, workers in these occupations saw their establishments become more concentrated in occupation during this time period. Using the Herfindahl measure of occupational concentration, the increase in occupational concentration are seen in the unadjusted trend lines for the three bottom quintiles, but after controlling for establishment characteristics and own-occupation, this trend persists only for the lowest-paid quintile of occupations. Using the predicted variance of ln(wages) based on occupations measure, only the bottom two quintiles of occupations saw an overall increase in occupational concentration (decreasing predicted variance), although all quintiles do see increased occupational concentration from 2014 - 2016. The between-occupations component of this measure similarly shows an overall increase in occupational concentration only for the lowest-wage quintile of occupations, with increased concentration in every group during 2014 – 2016.

Continuous-time versions of similar regressions are shown in Tables 5 (for all occupations) and 6 (by quintiles of occupation). These regressions take the form

$$(4) \qquad Occupational\ Concen_{jt} = \gamma Survey\ Date_t + \theta I(May\ Survey) + \delta X_{ijt} + \varepsilon$$

with γ coefficients scaled to measure changes in occupational concentration per decade of time, and X additional variables as above. The regressions in Table 5 show that the Herfindahl measure of occupational concentration shows an increase in occupational concentration over time, but occupations and employer characteristics explain about 95% of this increase. Meanwhile, the predicted variance of ln(wages) measure of occupational concentration (and its between-wages component) has risen over time, showing a trend of decreasing concentration overall for this measure. Table 6 shows the greatest increases in the Herfindahl measure of occupational concentration—after including controls for occupation and establishment characteristics—occur in the bottom and top quintiles of occupations. After including these controls, there are small but very significant decreases in the predicted variance of ln(wages) (increasing concentration over time) only for the occupations in the bottom quintile.

Both Figures 4 and 5 and Tables 4 and 5 measure changes in mean levels of these occupational concentration measures. Examining histograms[11] of the full distribution of these occupational concentration measures over time shows that the Herfindahl measure has had a fairly similar distribution over time, but the distribution of the predicted variance of ln(wages) has become increasingly bimodal, with modal values falling for establishments employing people in typically low-wage occupations and rising for employers of higher-wage occupations.

Figures 3 and 6 help to explain how the predicted-variance measures can be falling (increased concentration) for workers in typically low-wage occupations while the same measures can be rising (decreased concentration) for workers in typically low-wage occupations. Figure 3 showed the polarization of overall employment, with rising shares of employment in the top and bottom quintiles of occupations at the expense of the middle three quintiles of occupations. Figure 6 shows the fraction of workers in each quintile of occupation who work in establishments without any workers in other quintiles. It is unsurprising that workers in all other quintiles of occupation are growing less likely to have any coworkers in the middle three quintiles, as the middle quintile occupations have declining shares of overall employment over time. However, Figure 6 also shows that workers in the bottom three quintiles increasingly have no coworkers in the top quintile of occupation, and workers in the top three quintiles increasingly have no coworkers in the bottom quintile of occupation, although the bottom and top quintiles of the occupational distribution have increasing shares of overall employment over time.

To understand the impact of these trend in employment by occupational quintiles on the predicted variance of wages for establishments, consider three occupation groups: low-wage occupation group L (corresponding to the occupations in the lowest-paid quintile of occupations), middle-wage occupation group M (corresponding to the occupations in the middle

[11] Available upon request

three quintiles), and high-wage occupations H (corresponding to the occupation in the highest-paid quintile of occupations), with mean wages for occupations in each group $\overline{w_{OL}} < \overline{w_{OM}} < \overline{w_{OH}}$ and within-occupations wage variances by group $\sigma_{OL}^2 < \sigma_{OM}^2 < \sigma_{OH}^2$. Each establishment j contains $n_L \geq 0$ workers in the low-wage occupation group, $n_M \geq 0$ workers in the middle-wage occupation group, and $n_H \geq 0$ workers in the high-wage occupation group, with $n_L + n_M + n_H = n_J$. The predicted variance of wages for this establishment is $\widehat{V_J} = \frac{\sum_{occs\ \epsilon L} n_o\left[\left(\overline{w_{OL}} - \widehat{\overline{w_J}}\right)^2\right]}{n_j} + \frac{\sum_{occs\ \epsilon\ M} n_o\left[\left(\overline{w_{OM}} - \widehat{\overline{w_J}}\right)^2\right]}{n_j} + \frac{\sum_{occs\ \epsilon\ H} n_o\left[\left(\overline{w_{OH}} - \widehat{\overline{w_J}}\right)^2\right]}{n_j} + \frac{\sum_{occs\ \epsilon L} n_o {\sigma_{oL}}^2}{n_j} + \frac{\sum_{occs\ \epsilon M} n_o {\sigma_{oM}}^2}{n_j} + \frac{\sum_{occs\ \epsilon H} n_o {\sigma_{oH}}^2}{n_j}$. For workers in occupation group L, employing establishments have higher n_L, lower n_M, and lower n_H,[12] and, as shown in Figure 6, growing numbers of workers in occupation group L work in establishments with $n_M = n_H = 0$. There is little variation in wages between the occupations within the lowest-paid quintile, and so the changing distribution of occupations in their employment establishments reduces the typical value of $\left(\overline{w_{OL}} - \widehat{\overline{w_J}}\right)$, increases the weight on this component of the predicted wage variance and the ${\sigma_o}^2$component for low-wage (low-variance) occupations, and reduces the weight of the other 4 components of the predicted wage variance. This reduces both the between-occupations and within-occupations components of $\widehat{V_J}$, lowering $\widehat{V_J}$ for the workers in occupation group L. For workers in occupation group H, employing establishments have lower n_L, lower n_M, and higher n_H,[13] and, as shown in Figure 6, growing numbers of workers in occupation L work in establishments with $n_L = n_M = 0$. Although values of $\widehat{\overline{w_J}}$ will increase in these establishments, bringing these predicted means closer to the occupation-specific means for the occupations in occupation group H $\overline{w_{OH}}$, the wide variation in variation in wages between the occupations in this group means that a greater weight on this term of the predicted wage variance can increase the between-occupations component of $\widehat{V_J}$. Meanwhile, the high wage variance within the occupations of group H means that a greater weight on ${\sigma_{oH}}^2$ will also increase the within-occupations component of $\widehat{V_J}$ for the establishments employing H group occupations.

A closer look at Accountants, an example high-wage occupation from Table 2, shows how a high-paying occupation can have growing occupational concentration over time by the Herfindahl measure at the same time that is has shrinking occupational concentration over time by the predicted wage variance measures (increased variance). During this period, the share of Accountants working in the Professional services sector had no trend, employing 39% of accountants in both November 2002 and November 2016, the share working in the manufacturing sector fell from 10% to 8%, and the share working in the management of companies rose from 4% to 8%. Establishments that employed accountants became increasingly specialized during this period in 2 occupational categories: Health Diagnosing and Treating Practitioners (291) and Business Operations Specialists (131), with some growth in the number of Financial Specialists (132), the category that includes accountants. All of these are high-wage occupations. Meanwhile, employers of accountants employed fewer people in occupational categories such as Other Office and Administrative Support Workers (439), Production occupations (512, 514, 519, 537), Secretaries and Administrative assistants (436), and Financial Clerks (433). This changing occupational distribution, consistent with the occupational

[12] Exact numbers to confirm this statement are available on request.
[13] Exact numbers to confirm this statement are available on request.

polarization literature, meant rising predicted wage variances (lower concentration by variance measures), even as these employers concentrated employment in particular occupations (greater concentration by the Herfindahl measure).

Further heterogeneity in occupational concentration trends—by state-level unionization rates, establishment size, establishment age, industrial sector, and Employer Tax Identification Number (EIN) size, is described in Appendix B.

For low-wage workers, all measures show a clear trend of increased occupational concentration. The previous section showed that occupational concentration is strongly negatively associated with wages for workers in low-wage occupations. Thus, increasing concentration for low wage workers means a widening gap between their wages and the wages of other workers. Section IV explores what these patterns mean for wage inequality.

V. Occupational Concentration and Wage Inequality

The power of occupational concentration to explain wages as shown in Section IIb suggests that occupational concentration may explain some of the residual between establishment differences in wages highlighted by Barth, Bryson, Davis, and Freeman (2016). Moreover, Song, Price, Guvenen, and von Wachter (2016) show that the vast majority of pay-inequality growth at small and medium-sized firms 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. This section presents evidence showing that occupational concentration does indeed describe a substantial amount of wage variation across establishments, and that the amount described by these occupational concentration measures has grown over time. Furthermore, changes in occupational concentration between employers can explain a substantial amount of the overall growth in wage inequality during this period.

Va. Wage inequality between establishments

Table 7 documents the additional degree to which occupational concentration can account for between establishment ln(wage) variation. The table gives R^2 values from regressions of establishment mean ln wage levels on the occupational concentration measures developed here, the group of industry, establishment size, and geographic variables similar to those used by Barth et al. (2016), and combinations of the two. Looking across the columns of Table 7, all three occupational concentration measures explain a large and growing portion of ln wage variation between establishments. The Herfindahl measure of occupational concentration in the second row explains 11.7% of ln(wage) variation between establishments in November 2002 and 12.7% of ln(wage) variation between establishments in November 2016. The predicted variance of wages for the establishment based on the occupations employed, shown in the third row, explains 24.4% of ln(wage) variation between establishments in November 2002 and 35.1% in November 2016. The between-occupation component of this variance, shown in the fourth row, explains 8.8% of variation between establishments in November 2002 and 12.4% in

November 2016.[14] For comparison, the fifth row shows the amount of ln(wage) variation between establishments that can be explained by using the occupational composition of each establishment to predict its average ln(wage), which increases from 69.7% in November 2002 to 73.1% in November 2016.

In rows 6-8, estimates are given for models combining industry, size, and state variables from the baseline regression in row 1 with measures of occupational concentration. Adding the Herfindahl index (row 6) explains only about a percentage point more of the variation in establishment-level mean wages than the other characteristics given in row 1, but adding the predicted variance of wages for the establishment based on the occupations employed (row 7) increases the amount of explained wage variation by 3-4 percentage points in each year. Adding all the occupational concentration measures (row 10) increases the amount of ln(wage) variation between establishments that can be explained to 74.9% in November 2002 and 75.0% in November 2016. For comparison, combining the predicted ln(wage) for each establishment with other establishment characteristics (row 9) can explain 81.9% of ln(wage) variation between establishments in November 2002 and 82.6% in November 2016, an amount that is scarcely increased by adding the measures of occupational concentration (row 11).

The results in Table 7 show that occupational concentration substantially increases the amount of wage variation between establishments explainable by observable characteristics and that the amount of establishment-level wage variation explained by occupational concentration is slightly increasing over time. In data collected in November 2002, typically measured characteristics explained 70.0% of wage variation between establishments (row 1), and adding measures of occupational concentration (row 10) increased the amount by 5 percentage points. By 2016, adding measures of occupational concentration increased the total amount of establishment-level wage variation explained to 75.0% from 67.1%—an increase of 7.9 percentage points.

Vb. Overall wage inequality growth

As shown in Sections IIb and IIc, the workplaces for workers in the bottom quintile of occupations have become increasingly occupationally concentrated over time, and these are the workers whose wages are most adversely affected by occupational concentration. Meanwhile, the workplaces for workers in the top quintile of occupations have (by the predicted ln(wage) variance measure) become less occupationally concentrated over time, and these are the workers for whom wages increase with occupational concentration. This suggests that occupational concentration—particularly by the predicted variance of ln(wage) measure—may affect wage inequality at the individual-worker level. It is straightforward to calculate counterfactual wage distributions for 2016 using the method of DiNardo, Fortin, and Lemieux 1996 (DFL)[15] by

[14] A full set of R^2 values for establishment-level regressions and EIN-level regressions in every panel of the data is available upon request.

[15] The DiNardo, Fortin, and Lemieux (1996) methodology of creating counterfactual distributions for a later year if observable characteristics were held fixed at their distribution in an earlier year is to (1) combine the data for the earlier and later years and run a probit regression of the probability that an observation with a particular set of observable characteristics came from the earlier year and then (2) use the predicted values from this probit regression to create new weights for each observation in the later year.

reweighting observable characteristics to their distributions in November 2002/May 2003. Thus, portions of increased wage inequality growth from 2002/2003 to 2016 can be attributed to changes in the distribution of occupations, employment by industry, state, employer size, occupational concentration, and combinations of these factors by holding subsets of the characteristics to their 2002/2003 distributions.[16]

Table 8 shows the results of DFL-type reweightings for the observable characteristics of detailed industry (at the 4-digit NAICS level), state, employer size, occupation (at the 3-digit SOC code level), and measures of occupational concentration, and combinations of these variables. The overall variance of ln wages increased from .369 in November 2002/May 2003 to .386 in May/November 2016, and all of this increase is due to between-establishment wage variance increasing from .212 to .233. For each reweighting, rows of table 8 show the resulting variance of ln(wages) and the percentage of the growth in the overall, between-establishments and within-establishment ln(wage) variances that can be explained by changes in the distribution of the observable characteristic(s). In addition, each row shows real wages in $2000 for selected percentiles[17] of the May/November 2016 wage distribution and the 50-10, 90-50 and 90-10 wage range. Row (1) of Table 8 gives the levels of overall, between-establishment, and within-establishment wage variance observed in November 2015 without any reweighting. Further rows of the table show results of reweightings for single characteristics or for selected combinations of observable characteristics. The DFL-type reweightings were performed for all combinations of characteristics, and the particular results shown in the table are chosen to summarize these results.

The first set of reweighting characteristics examined in Table 8 are the variables used as controls in the regressions shown in Tables 3-6. Row (2) shows that changes in 4-digit industry can account for a substantial amount of overall wage variance growth from November 2002/May 2003 to May/November 2016. Had the distribution of employment by 4-digit industry remained at the 2002/2003 levels, overall wage variance would only have increased to .3788 instead of to .3861, as shown by comparing row (2) to row (1). Changes in the distribution of this variable explain 42% of the growth in overall ln wage variance and 31% of the growth in between-establishment ln wage variance from 2002/2003 to 2016. Similarly, reweighting observations in 2016 to the earlier distribution of 3-digit occupation categories (row 5) explains 104% of the growth in overall ln wage variance and 65% of ln wage variance growth between establishments. Changes in the distributions of employment by establishment size classes (row 4) also explains some of overall ln wage variance growth, while by changes in the distribution of employment by states (row 3) does not explain any of this ln(wage) variance growth.

Because the occupational concentration measures are continuous rather than categorical variables, these variables are divided into deciles for this reweighting exercise. Because Figure 5 and Table 6 show that occupational concentration is changing in different ways for different

[16] While Fortin, Lemieux, and Firpo (2011) advocate that this method should be replaced whenever possible by RIF-regression methods, which can also examine the impact of changes in the returns to characteristics and not just changes in composition. Preliminary research using the data from Handwerker and Spletzer (2014) shows that the interval nature of the OES data distorts the results of RIF regressions, and so work presented here relies instead on the older method of Dinardo, Fortin, and Lemiuex (1996).

[17] Since the OES survey collects employment by wage interval, not exact wages, these percentiles are calculated assuming a uniform distribution of wages within each interval.

quintiles of occupations, the occupational concentration variables are interacted with quintiles of occupation.[18] Row (7) shows that changes in the distribution of the Herfindahl measure of occupational concentration interacted with quintiles of the occupational distribution more than explains the growth in overall ln wage variance growth from 2002/2003 to 2016—126%, as well as 78% of the growth of wage variance between establishments. However, much of this explanatory power comes from this variable's interaction with the changing distribution of the quintiles of the wage distribution, illustrated in Figure 3. Row (6) shows that the changing distribution of occupations by quintile, rather than by the 3-digit occupations of row (5), explains 123% of the overall growth the variance of ln(wages), as well as 71% of the growth in the variance between establishments. Row (8) shows that changes in the distribution of the predicted variance of ln(wages) of establishments, interacted with quintiles of occupation, explains even more of the growth in overall ln wage variance growth from 2002/2003 to 2016—144%, as well as 93% of the growth of wage variance between establishments. The between-occupations portion of the predicted ln(wage) variance of establishments, interacted with quintiles of occupation, explains 127% of the growth in wage variance, as shown in row (9).

Further rows of Table 8 show reweighting for selected combinations of observable characteristics. Without including any of the occupation concentration measures, changes in the distributions across establishment size categories and occupational quintiles (row 12) are able to explain more of the rise in overall wage variance (124%) and the rise in between-establishment wave variance (73%) than any other combination of these establishment characteristics. The amount that changes in occupational concentration contribute to the growth in wage variance is clear when comparing row (12) with rows (8) and row (11). Row 11 includes industry, Herfindahl index categories, and predicted variance of ln(wage categories, with occupational quintiles included as interactions. This combination of variables gives the largest amount of between-establishment wage variance growth explained (95%). Row (8) gives an increase of 20 percentage points of total wage variance growth and row (11) gives an increase of 22 percentage points of wage variance growth between establishments above the amount explained without the occupational concentration variables in row (12). Using additional reweighting variables does not always explain more of the growth in variance; combining all available variables, as in row (10), can only explain 85% of overall wage variance growth (although it can explain 79% of wage variance growth between establishments).

The contrast between rows (8) and (9) of Table 8 explains a great deal about the impact of different forms of occupational concentration on the wage distribution. The interaction of the predicted variance of ln(wages) of establishments with quintiles of occupation explains much more of the growth in overall wage variance than do the quintiles of occupation alone, while the interaction of only the between-occupations portion of this predicted variance with quintiles of occupation explains barely any more of the growth in overall wage variance than quintiles of occupation alone. How do these variables differ? The between-occupations portion of the predicted variance in wages for each establishment excludes the within-occupation component of predicted variance, the portion that is strongly correlated with employing high-wage occupations in the establishment. Establishments employing low and middle wage occupations may have very similar values of the between-occupations portion of the predicted variance of wages when

[18] This follows the example of Goldschmidt and Schmieder in section V.C. of their paper, who use indicators for deciles of the firm wage effect multiplied by dummies for the frequently outsourced occupations.

compared with establishments employing middle and high wage occupations, but there will be a much greater total predicted variance of wages for the establishments employing middle and high wage occupations. To check that employment of high-wage occupations in the establishment matters for changes in the wage distribution, row (13) of Table 8 shows the impact of reweighting observations in 2016 to the 2002/2003 distribution of a dummy variable for the establishment having any workers in occupations in the top quintile, interacted with the quintiles of occupation (for workers in the top quintile, this is instead a dummy variable for their establishment employing exclusively workers in top quintile occupations). This simple dummy variable interaction explains slightly more of overall variance growth—and more of the variance growth between occupations—than changes in the distribution of the occupational quintiles alone. Adding this new measure to combinations of other reweighting variables, row (14) shows it is possible to explain as much as 99% of the growth in ln(wages) between establishments.

The discussion of reweighting results thus far has focused entirely on the impact of reweighting observable characteristics on overall and between-establishment ln(wage) variance. However, these reweightings also show how changes in observable characteristics impact other measures of the distribution of wages. For example, while occupational concentration measures clearly matter in explaining the increase in wage inequality, these measures might not affect the variance of wages symmetrically; wages could increase or decrease only on one side of the distribution. Columns at the right side of Table 8 show real wages in $2000 for selected percentiles[19] of the May & November 2016 wage distribution and the associated 50-10, 90-50 and 90-10 wage gaps for these reweightings. Actual percentiles for the 2016 wage distribution are given in row (1).

Reweighting the 2016 data to the 2002/2003 distribution of each measure of occupational concentration interacted with occupational quintiles (rows 7, 8, 9, and 13) would raise wages at the 10th percentile and lower them at the 90th percentile—reducing the spread in the wage distribution. However, while all of these occupational concentration reweightings would raise hourly wages at the 10th percentile by a few cents more than reweighting by occupational quintiles alone (row 6), only the reweighting by the top-quintile employment dummies (row 15) would lower wages at the 90th percentile by more than reweighting by occupational quintiles alone. The combination of reweighting characteristics in rows (11) and (17) bring the greatest reduction in the 90-50 and the 90-10 wage gaps while leaving the 50-10 gap little changed. Changes in the in distribution of occupational concentration measures between 2002/2003 and 2016 have a large impact on wage inequality, and this impact is mostly manifested in the upper part of the wage distribution.

VI: Occupational Concentration: Establishment or Firm Measures?

Song, Price, Guvenen, Bloom, and von Wachter (2016) argue that the unit of importance for wage inequality should be the firm and not the establishment. In thinking about occupational concentration, 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

[19] Since the OES survey collects employment by wage interval, not exact wages, these percentiles are calculated assuming a uniform distribution of wages within each interval.

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 concentration at the establishment level because establishments are the sampling units of the OES, and the OES 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 OES 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 there is no way to link together all of the establishments in these data for very large firms without a tremendous amount of manual review. Thus, while it is straightforward to recalculate measures of occupational concentration at the EIN level and repeat the analyses above, such EIN-level measures are not true firm-level measures.

Using EIN-level measures of occupational concentration instead of establishment-level measures has remarkably little impact on any of the main results in this paper.[20] The relationship between EIN-level measures of occupational concentration and wages is very similar to that shown for establishment-level measures in Tables 3 and 4. The main difference is that in regressions of equation (2), for both the Herfindahl and predicted variance of ln(wage) measures of occupational concentration, the coefficients α indicate a significantly larger impact of occupational concentration on wages for the EIN-level measures than for establishment-level measures. Trends in EIN-level measures of occupational concentration over time are very similar to those for establishment-level measures in Tables 5 and 6. However, the increased concentration of workers in the bottom quintile is significantly larger when measured with establishment-level measures of occupational concentration than when using the EIN-level equivalents.

Reweighting the May/November 2016 data to the November 2002/May 2003 distribution of EIN-level measures of occupational concentration also yields very similar results to those shown in Table 8. The explanatory power of the predicted variance of wages based on the EIN-level distribution of occupations interacted with quintiles of occupation is even greater than the establishment level version of this measure, explaining 150% of the growth in ln wage variance, rather than the 144% shown in row (8) of table 8. Using this measure as well as the Herfindahl

[20] Note that 92% of the EINs in the OES data (containing 64% of weighted employment) are associated with only one establishment per panel in which they appear. However, trends in occupational concentration are very similar for employers with 10 or more establishments.

interaction with occupation quintile and the dummy variable for the presence of top-quintile occupations in an EIN yields a reweighting that can explain 161% of overall ln wage variance growth, and 104% of the growth in ln wage variance between EINs, rather than the 99% shown for establishments in row (14) of Table 8. Without these occupational concentration variables, combinations of other reweighting variables explain no more than 123% of overall wage inequality growth and 62% of the growth in wage variance between EINs—a difference of 38 percentage points for overall wage inequality growth and 42 percentage points for wage inequality growth between EINs.

VII. 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 in particular occupations and industries, this paper is among the first to connect the two with a study of the impact of the changing distribution of occupations between employers on wage inequality in the United States. Occupational concentration is a description of outsourcing, and this paper uses multiple measures of occupational concentration (at both the establishment and employer tax-ID levels) to examine the impact of outsourcing on wages and on wage inequality. These measures show greater occupational concentration for the occupations used to study 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 outsourced sector. For example, occupational concentration is higher for janitors when there are employed in establishments in the janitorial services industry.

The advantage of measuring outsourcing with occupational concentration is that these measures can be calculated for every employee of every employer, not only for "case study" occupations. This paper shows that occupational concentration is strongly and significantly related to wages across all occupations, and has a particularly strong negative wage impact for the quintile of workers in the typically lowest paid occupations, even after controlling for the occupations of employees and the various observable characteristics of their employers. (However, occupational concentration has a positive impact on wages within employers of 10,000 or more employees). By every measure of occupational concentration—and for employers of every size—there is an increase in concentration over time for the quintile of workers in the lowest-wage occupations. Employment polarization means fewer coworkers in middle-wage occupations, but low-wage workers also have fewer coworkers over time in high-wage occupations, even as low-wage and high-wage occupations make up a growing share of employment. The pattern of time trends across measures of occupational concentration, with and without controlling for employer characteristics, is consistent with the idea that in the economy as a whole, companies are "de-verticalizing" by outsourcing functions not integral to employers' missions, particularly if these outsourced tasks are done by workers paid lower wages than the "core workers" in the establishment.

Song, Price, Guvenen, and von Wachter (2016) 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 higher-paid workers—to firms that pay lower wages. Occupation is a just such a characteristic affecting workers' wages, and this paper shows that workers in low-wage occupations are increasingly concentrated at employers with fewer high-wage occupations, contributing to wage inequality growth.

Using detailed data about the occupational composition of each employer, the distribution of occupations and occupational concentration, combined with other employer characteristics, can explain 82% of the variation in establishment-level ln(wages) in November 2002, and 83% of the variation in establishment-level ln(wages) in November 2016. A rising share of this between-establishment wage variation is explained by the distribution of occupations and occupational concentration, while a falling share can be explained by other employer characteristics. The changing distributions of occupations and occupational concentration can also more than explain the growth in ln(wage) variance during this time period. Changes in the distribution of establishments by various measures of occupational concentration along with other observable variables, can explain as much as 144% of measured ln(wage) variance growth at the individual level and 99% of the measured increase in wage variance between establishments. These data are best suited to examine occupational concentration at the establishment level, but when larger measures of employers (EINs) are used instead, very similar patterns appear, and even more of the growth in wage inequality can be explained. The specific measures of occupational concentration with the greatest power to explain wage inequality growth involve the presence or absence of high-wage occupations at the employer. Thus, the growing physical separation of workers in low-wage occupations from workers in high-wage occupations appears to be an important part of wage inequality growth.

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 (2009). "New Evidence on the Returns to Job Skills." *American Economic Review Papers and Proceedings*, Vol. 99, No. 2, pp. 52-57.

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., and David Dorn (2013). "The growth of low-skill service jobs and the polarization of the US labor market." *The American Economic Review,* Vol.103, No. 5, pp. 1553-1597.

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.

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, and Elizabeth Weber Handwerker (2016). "Longitudinal data from the Occupational Employment 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.

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.

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 (2015). "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure," NBER Working Paper #21366. *The Quarterly Journal of Economics, forthcoming?*

Goos, Maarten and Alan Manning (2007). "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *Review of Economics and Statistics*, Vol. 89, No. 1, pp. 118-133.

Goos, Maarten, Alan Manning and Anna Salomons (2009). "Job Polarization in Europe," *American Economic Review Papers and Proceedings*, Vol. 99, No. 2, pp. 58-63.

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.

Lemieux, Thomas (2008). “The Changing Nature of Wage Inequality.” *Journal of Population Economics*, Vol. 21, No. 1, pp. 21-48.

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

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. (2016) Firming Up Inequality. No. w21199. National Bureau of Economic Research, 2015.

Data Appendix

This paper uses Occupational Employment Statistics (OES) Survey microdata. The OES 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. This data collection occurred in October, November, and December, until 2001; since November 2002, data has been collected for about 200,000 establishments each November and another 200,000 each May. As described in Dey and Handwerker, the OES uses a complex sample design intended to minimize the variance of wage estimates for each occupation within industries and geographic areas. Thus, establishments expected to employ occupations with greater variation in wages have relatively larger probabilities of selection and lower estimation weights.

The OES survey form is a matrix of detailed occupations and wage intervals. For large establishments, the survey form lists 50 to 225 detailed occupations; these occupations pre-printed on the survey form are selected based on the industry and the size of the establishment. Small establishments write descriptions of the work done by their employees, which are coded into occupations by staff in state labor agencies. Wage intervals on the OES 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 OES program obtains the mean of each wage interval every year from the National Compensation Survey (NCS). These mean wages are then assigned to all employees in that wage interval. The OES survey is not designed to produce time series statistics. Time series in this paper are produced using the methodology described in Abraham and Spletzer (2010) to reweight the data to November or May benchmarks of total employment by detailed industry and by broad industry and establishment size groups from the Quarterly Census of Employment and Wages (QCEW).

The OES began collecting data using the Standard Occupational Classification System in 1999, and had a change of industry classification systems from SIC to NAICS (2002) soon thereafter. Beginning with the 2002 OES survey, establishments were classified by 6 digit NAICS (2002) codes, and the OES staff recoded much of the 1999, 2000 and 2001 OES microdata to use these same NAICS codes. The analyses in this paper begin with the OES microdata from 1999 in order to be able to use consistent industry and occupation classifications throughout. Certain SOC and NAICS codes are combined to make groups consistent across the 2007 and 2012 NAICS revisions and the 2010 revision to the SOC. Nonetheless, as noted in Abraham and Spletzer, there are some inconsistencies of SOC coding in the initial years that the OES program used this coding system. In particular, the previous occupation coding system allowed workers with titles such as “financial manager” or “marketing manager” to be classified as “managers” even if they did not supervise any employees, while the SOC system would not consider such titles to be managerial occupations. This particularly affected the occupation coding of small establishments by state workforce agencies in 1999 and 2000 (and to a lesser extent in 2001), before all state-level staff had received training about this aspect of SOC classification; large establishments were not subject to occupational coding by state-level staff and were not affected by this inconsistency in coding procedures.

The OES *cannot* measure inequality in the top percentiles of the wage distribution. Earnings of individuals at the very top of the wage distribution are topcoded in the OES—the uppermost interval in the recent OES surveys is "$208,000 and over" (interval ranges vary by year). Averaged across all years, the uppermost interval contains roughly 1.3 percent of employment. Handwerker and Spletzer 2014 compare wage data in the OES with wage data from the outgoing rotation groups of the CPS, and have two main findings. First, the interval nature of wage collection in the OES has almost no impact on overall wage variance trends. Second, reweighted OES data broadly replicate basic CPS wage distribution trends, beginning in 1998. Overall wage distributions in each year are similar, as well as overall variance trends, variance trends by sector, industry groups, and occupation groups. In both the OES and the CPS, industry groups alone explain 15-17% of wage variation, although industry groups explain slightly more of the variation in the (employer-reported) OES than in the (employee-reported) CPS. Occupational groups alone explain more of the variation in wages in the OES (about 40%) than these same variables explain in the CPS (about 30%). The amount of wage variance explained by occupation is also growing more quickly in the OES than in the CPS.

Handwerker and Spletzer 2016, examine the decomposition of total wage variance in the OES into its within-establishment and between establishment components at length. Updating those findings, over the period of Fall 1999 through November 2015, 60% of wage variance is between establishments, while 98% of the growth in overall wage variance over this period is between establishments. Handwerker and Spletzer 2016 also find that similar amounts of establishment-level wage variance in the OES can be explained by broad industry groups to the amount found by Barth, Bryson, Davis, and Freeman. However, more of the establishment-level wage variance can be explained by detailed industry in the OES data than in the Census data, echoing findings comparing OES and CPS data.

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	1.94	1.1%	1
353	Food and Beverage Serving Workers	1.94	6.5%	1
393	Entertainment Attendants and Related Workers	1.98	6.9%	1
352	Cooks and Food Preparation Workers	2.05	9.3%	1
392	Animal Care and Service Workers	2.09	9.4%	1
399	Other Personal Care and Service Workers	2.09	11.1%	1
412	Retail Sales Workers	2.10	18.6%	1
372	Building Cleaning and Pest Control Workers	2.10	20.9%	1
452	Agricultural Workers	2.11	21.0%	2
536	Other Transportation Workers	2.13	21.2%	2
516	Textile Apparel and Furnishings Workers	2.15	21.9%	2
311	Nursing Psychiatric and Home Health Aides	2.16	23.8%	2
396	Baggage Porters Bellhops and Concierges	2.19	23.9%	2
395	Personal Appearance Workers	2.20	24.3%	2
373	Grounds Maintenance Workers	2.21	25.0%	2
339	Other Protective Service Workers	2.23	26.0%	2
397	Tour and Travel Guides	2.23	26.0%	2
513	Food Processing Workers	2.24	26.7%	2
537	Material Moving Workers	2.24	30.6%	2
379	Other Buildings, Grounds, and Maintenance Occupations	2.27	30.6%	2
259	Other Education Training and Library Occupations	2.28	30.9%	2
435	Material Recording Scheduling Dispatching and	2.29	33.8%	2
432	Communications Equipment Operators	2.29	34.0%	2
473	Helpers Construction Trades	2.31	34.2%	2
459	Other Farming, Fishing, and Forestry Occupations	2.32	34.2%	2
453	Fishing and Hunting Workers	2.33	34.2%	2
439	Other Office and Administrative Support Workers	2.35	37.3%	2
517	Woodworkers	2.36	37.5%	2
512	Assemblers and Fabricators	2.40	39.2%	2
434	Information and Record Clerks	2.40	43.5%	3
519	Other Production Occupations	2.40	45.9%	3
319	Other Healthcare Support Occupations	2.42	47.0%	3
351	Supervisors of Food Preparation and Serving Workers	2.45	47.7%	3
433	Financial Clerks	2.48	50.5%	3
332	Fire Fighting and Prevention Workers	2.48	50.5%	3
533	Motor Vehicle Operators	2.49	53.5%	3
454	Forest Conservation and Logging Workers	2.50	53.5%	3
515	Printing Workers	2.53	53.8%	3
219	Other Community and Social Service Occupations	2.53	53.8%	3

3-digit SOC code	*SOC Title*	*Average ln(wage)*	*Cummulative percentage of employment*	*Occupation Quintile*
252	Preschool Primary Secondary and Special Education	2.54	54.4%	3
514	Metal Workers and Plastic Workers	2.56	56.2%	3
394	Funeral Service Workers	2.56	56.3%	3
436	Secretaries and Administrative Assistants	2.57	59.0%	3
419	Other Sales and Related Workers	2.57	59.8%	4
253	Other Teachers and Instructors	2.58	60.1%	4
333	Law Enforcement Workers	2.58	60.1%	4
391	Supervisors of Personal Care and Service Workers	2.59	60.2%	4
211	Counselors Social Workers and Other Community and	2.59	61.2%	4
371	Supervisors of Building and Grounds Cleaning and	2.62	61.4%	4
493	Vehicle and Mobile Equipment Mechanics Installers	2.63	62.6%	4
312	Occupational Therapy and Physical Therapist Assistants	2.65	62.8%	4
499	Other Installation Maintenance and Repair Occupations	2.65	64.9%	4
212	Religious Workers	2.66	64.9%	4
534	Rail Transportation Workers	2.67	65.0%	4
475	Extraction Workers	2.68	65.1%	4
474	Other Construction and Related Workers	2.70	65.3%	4
292	Health Technologists and Technicians	2.71	67.4%	4
472	Construction Trades Workers	2.71	71.1%	4
274	Media and Communication Equipment Workers	2.72	71.2%	4
271	Art and Design Workers	2.77	71.7%	4
331	Supervisors of Protective Service Workers	2.77	71.7%	4
194	Life Physical and Social Science Technicians	2.77	71.9%	4
411	Supervisors of Sales Workers	2.78	73.2%	4
254	Librarians Curators and Archivists	2.79	73.2%	4
272	Entertainers and Performers Sports and Related	2.80	73.6%	4
451	Supervisors of Farming Fishing and Forestry Workers	2.81	73.6%	4
492	Electrical and Electronic Equipment Mechanics Install	2.81	74.1%	4
232	Legal Support Workers	2.83	74.4%	4
531	Supervisors of Transportation and Material Moving	2.87	74.7%	4
239	Other Legal Occupations	2.88	74.7%	4
431	Supervisors of Office and Administrative Support	2.90	75.8%	4
535	Water Transportation Workers	2.90	75.8%	4
299	Other Healthcare Practitioners and Technical Occs	2.91	75.9%	4
173	Drafters Engineering Technicians and Mapping	2.91	76.5%	4
273	Media and Communication Workers	2.93	77.0%	4
511	Supervisors of Production Workers	2.98	77.5%	4
413	Sales Representatives Services	3.01	78.9%	4
518	Plant and System Operators	3.01	79.0%	4
153	All other Computer and Math Occupations	3.05	79.0%	4

3-digit SOC code	SOC Title	Average ln(wage)	Cummulative percentage of employment	Occupation Quintile
414	Sales Representatives Wholesale and Manufacturing	3.08	80.7%	4
491	Supervisors of Installation Maintenance and Repair	3.08	81.0%	5
131	Business Operations Specialists	3.11	83.8%	5
471	Supervisors of Construction and Extraction Workers	3.11	84.2%	5
193	Social Scientists and Related Workers	3.16	84.3%	5
171	Architects Surveyors and Cartographers	3.17	84.4%	5
132	Financial Specialists	3.17	86.3%	5
251	Postsecondary Teachers	3.18	86.8%	5
159	Computer and Math Occupations, NEC	3.24	86.8%	5
532	Air Transportation Workers	3.26	87.0%	5
192	Physical Scientists	3.31	87.2%	5
151	Computer Specialists	3.32	89.8%	5
191	Life Scientists	3.32	90.0%	5
119	Other Management Occupations	3.33	91.3%	5
152	Mathematical Science Occupations	3.35	91.4%	5
291	Health Diagnosing and Treating Practitioners	3.39	94.8%	5
172	Engineers	3.43	96.0%	5
113	Operations Specialties Managers	3.60	97.2%	5
111	Top Executives	3.64	99.0%	5
112	Advertising Marketing Promotions Public Relations &	3.65	99.6%	5
231	Lawyers Judges and Related Workers	3.76	100.0%	5

Appendix B: Heterogeneity of Results

B1: Heterogeneity by state-level unionization rates

One factor which may impact both wages and the organization of production in terms of the variety of occupations at a workplace is unionization. The OES does not collect information on unionization patterns by employer, but it includes location of each establishment, and unionization rates vary strongly by state. Thus, state-level unionization rates are used to group the data into 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.

Following equation (2), the relationships between occupational concentration and wages are estimated separately for each unionization group of states. Across all occupations, the relationships α between occupational concentration (by all measures) and wages is significantly greater in the more highly unionized states. However, this reverses when establishment characteristics and occupational controls are included in equation (2), and after including these controls, the relationship between occupational concentration and wages (α) is significantly greater in the less unionized states than in the more highly unionized states. For workers in the lowest-paid quintile of occupations, occupational concentration (by all measures) matters more for wages in less unionized states both with and without controlling for establishment characteristics and occupation. Across all workers, time trends interacted with occupational concentration (β) have varying signs across measures of occupational concentration and the inclusion of controls, but for the workers in the lowest-paid quintile of occupations, these interactions are always significantly lower in less unionized states, indicating that the relationships between occupational concentration and wages for these workers are converging over time between the different groups of states.

Overall, workers in states with higher unionization levels work in slightly (but statistically significantly) less occupationally concentrated establishments. However, for workers in the lowest-paid quintile of occupations, this reverses; these workers have slightly higher occupational concentration in establishments located in states with higher unionization levels.

Differences in occupational concentration trends between less and more unionized states show that establishments are growing more concentrated over time in the less-unionized states, relative to the highly unionized states, by every measure. Following equation (4), occupational concentration as measured by the Herfindahl index appears to be increasing slightly (but significantly) faster in less-unionized states than in highly unionized states, both across all occupations and for occupations in the bottom quintile, and so workers in less unionized states now work in establishments with higher Herfindahl indices, on average, than workers in more unionized states. Occupational concentration as measured by the predicted variance of ln(wages) based on the occupational composition of establishments, and its between-occupations component also show less of an increase (more concentration) in the less unionized states, and after employer characteristics and occupation controls are included in equation (4), all measures of occupational concentration show statistically significant trends of increasing concentration in less unionized states. Differences in trends are similar for the lowest-paid quintile of occupations.[21]

B2: Heterogeneity by establishment age

Because the Occupational Establishment Survey data is sampled from the records of the BLS Quarterly Census of Employment and Wages, which BLS assembles into a longitudinal database of establishments, it is straightforward to link these datasets together and find a "birth date"—the first quarter with employment greater than zero—for each establishment. Dividing establishments into those born before the fourth quarter of 2002 (87% of all establishments observed) and those born afterwards (13%), the analyses above can be repeated separately for "old" and "young" establishments.

Young establishments are, in employment-weighted averages, more concentrated in occupations than old establishments, with higher Herfindahl indices of occupational concentration (.511 for young establishments compared with .386 for old establishments), and lower predicted variances of wages based on the occupational composition of establishments (.241 for young establishments compared with .276 for old establishments), with much of the difference due to the between-occupations component of this variance (.081 for young establishments compared with .111 for old establishments). This pattern is echoed at higher concentration levels for the establishments of workers in the lowest-paid quintile of occupations—for these workers as well, working in younger establishments means working in establishments more concentrated in occupation, by every measure.

Examining the relationships between occupational concentration and wages by establishment age, occupational concentration matters more for wages in old establishments than in young establishments than in young establishments. The difference between old and young establishments becomes much smaller after additional controls are added to equation 2, but occupational concentration—by every measure—still matters significantly more in old establishments than in young establishments. This is true across all workers as well as for workers in the lowest-paid quintile of occupations.

There are no clear patterns in differences in trends in occupational concentration by establishment age—differences vary greatly by which measure of occupational concentration is used, whether controls for establishment characteristics are included, and which groups of occupations are examined. [22]

B3: Heterogeneity by establishment size

Subdividing establishments into the same size classes used in the regression controls (1-4 employees, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, and 500+), there is a mostly monotonic decrease in employment-weighted means of occupational concentration with establishment size,

[21] Results available upon request.
[22] Results available upon request.

by every measure of occupational concentration. Larger establishments are more heterogeneous in occupations. There is no clear pattern across establishment sizes of relationships between occupational concentration and wages for all workers. However, for the lowest-paid quintile of occupations, the strongest relationships between occupational concentration and wages—across all measures, with and without controlling for other observable characteristics—appears in middle-sized establishments: those with 50-99 employees. [23]

B4: Heterogeneity by industrial sector

Dividing establishments into industrial sectors, establishments in the Construction (23), Administrative and Support and Waste Management (56), Accommodation and Food Services (72), and Other Services (except Public Administration) (81) have high levels of occupational concentration, by all measures. Management of Companies and Enterprises (55) has particularly low levels of occupational concentration, by all measures, while Utilities (22) has particularly low occupational concentration by the Herfindahl measure, private-sector Educational Services (61) has particularly low occupational concentration by the total predicated variance of wages based on the occupational distribution, and Health Care and Social Assistance (62) has particularly low concentration by the between-occupations portion of this predicted variance.

These sectors differ greatly in the fraction of their employment in occupations at different places in the wage distribution, but some of the same sectors stand out for having low or high levels of concentration among employers of workers in occupations in the lowest-paid quintile of the labor force. Examining occupational concentration levels by sector for workers in these low-paid occupations, again, Administrative and Support and Waste Management (56) and Other Services (except Public Administration) (81) have particularly high levels of occupational concentration, by all measures and Management of Companies and Enterprises (55) has particularly low levels of occupational concentration, while private-sector Educational Services (61) has particularly low occupational concentration by the total predicated variance of wages based on the occupational distribution. However, other sectors with particularly high or low levels of occupational concentration are different for workers in the bottom-paid quintile of occupations. For these workers, establishments in the Educational Services (61) sector also have particularly low levels of occupational concentration by all measures. Establishments in the Finance and Insurance Sector (72) have a high level of occupational concentration by the Herfindahl measure, while the Accommodation and Food Services (72) and Other Services (42) sectors have particularly high levels of occupational concentration by both predicted variance measures.

The relationship between occupational concentration and wages varies tremendously by sector, even after controlling for the occupations employed within each sector. After controlling for occupations, detailed industries, state, and establishment size, by every measure, greater occupational concentration is associated with lower wages within the Construction (23), Retail Trade (44), Transportation and Warehousing (48), Real Estate and Rental and Leasing (53), Professional, Scientific, and Tech Services (54), Management of Companies and Enterprises (55), Administrative and Support and Waste Management (56), Health Care and Social Assistance (62), Arts, Entertainment, and Recreation (71), Accommodation and Food Services

[23] Results available upon request.

(72), and Other Services (81) sectors. These relationships are particularly strong in the Transportation and Warehousing (48) sector. Other sectors have overall relationships between occupational concentration and wages that vary by measure of occupational concentration. For bottom-quintile workers, after controlling for observable characteristics, greater occupational concentration—by every measure—is associated with lower wages within all of the above sectors, as well as the Wholesale Trade (42) and (private-sector) Educational Services (61) sectors. For typically low-wage workers, the relationship between occupational concentration and wages is particularly strong across all measures of occupational concentration within the Utilities (22) and Arts, Entertainment, and Recreation (71) sectors.

Trends over time in occupational concentration also vary greatly by sector. Across all occupations and all measures of occupational concentration, after controlling for observable characteristics, occupational concentration is increasing within the Accommodation and Food Services (72) and Transportation and Warehousing (48) sectors. It is decreasing within the Construction (23), (private-sector) Educational Services (61), Manufacturing (31-33), Finance and Insurance (52), Real Estate and Rental and Leasing (53), and Other Services (81) sectors, with trends in the remaining sectors that vary in direction by measure of occupational concentration. For the lowest-paid quintile of occupations, after controlling for observable characteristics, occupational concentration is increasing by all measures within the Construction (23), Administrative and Support and Waste Management (56), Arts, Entertainment, and Recreation (71), and Accommodation and Food Services (72) sectors, while it is decreasing by all measures only within the Management of Companies and Enterprises (55) and (private-sector) Educational Services (61) sectors. [24]

B5: Heterogeneity by Employer tax Identification Number (EIN) size

Song et al find very different patterns of inequality growth from 1978 to 2013 for very large firms—those with Employer tax Identification Numbers (EINs) with 10,000 or more employees—than for smaller firms. They find that for smaller firms, nearly all inequality growth is between firms, explained by greater sorting of workers with higher worker fixed effects to firms with higher firm fixed effects, while very larger firms see nearly half of inequality growth happening within firms, with falling wages for their lowest-paid workers and rising wages for their highest-paid workers. The focus of this analysis is the establishment (except in section V) because establishments are the sampling units of the OES, and the OES 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 OES microdata can be linked with the EIN numbers that these establishments submit to the unemployment insurance system, as compiled by the Quarterly Census of Employment and Wages (QCEW), and the full employment level of each EIN can be calculated in each time period using the QCEW data. As discussed in Handwerker and Mason (2013), very large firms may use multiple EINs in the unemployment insurance system, and there is no way to link together all of the establishments in these data for very large firms without a tremendous amount of manual review. It is straightforward to group establishments into those that are part of very large EINs (those with 10,000 or more employees) and those that are not part of these large

[24] Results available upon request.

EINs, but the reader should be aware that many establishments not part of large EINs are nonetheless part of large firms that use smaller EINs in their quarterly reports to the unemployment insurance system.[25]

The reader should also be aware that very large establishments are included in the OES sample with certainty every 6 panels,[26] and the reweighting used to break apart the 6-panel groups of OES waves used for official OES publications into panel-specific microdata results in big swings of estimates (of any variable of interest) from one panel to the next for very large employers. This makes it difficult to measure trends in any variable for these employers. Nonetheless, the OES 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 OES data. However, there is likely a great deal of overlap between workers in typically-low-wage occupations and workers with low fixed effects.

Overall, the establishments of very large employers have lower Herfindahl values (.36, weighted by employment) of occupational concentration than the establishments of smaller employers (.42), and this is especially true for workers in the bottom quintile of occupations (.44 for very large employers and .50 for smaller employers). However, the predicted variance of wages based on occupational distributions is greater for the establishments of smaller employers (.27) than for very large employers (.26), although this reverses for workers in the bottom quintile (.20 for smaller employers and .21 for very large employers). The establishments of very large employers are much less likely than the establishments of smaller employers to have no workers in the top quintile of occupations (13% of the employees of very large employers have no top-quintile co-workers, compared with 25% of the employees of smaller establishments), and this is particularly true for workers in the bottom quintile (25% of the bottom-quintile employees of very large employers have no top-quintile co-workers, compared with 46% of the bottom-quintile employees of smaller employers). Relative to smaller employers, very large employers have establishments that are more diverse in the groups of occupations they employ, but are more concentrated in the wage range of the occupations they employ—but the very large employers of the workers of the bottom quintile of occupations workers are a little less concentrated in the wage range of occupations they employ than smaller employers of these low-paid occupations.

Overall, occupational concentration matters much more for wages within huge employers, by every measure of occupational concentration. However, this difference reverses once controls

[25] Song et al estimate that 23% of workers are in these very large firms; in the OES data, only 16% of workers are in such large firms. Various BLS projects have attempted to find all the EINs associated with particular sets of firms in particular time periods. Using all of the available links from these projects to group EINs together, it is possible to consider 20% of workers as associated with firms employing 10,000 or more; results using these larger groupings are little changed from the EIN-level results discussed in this section.

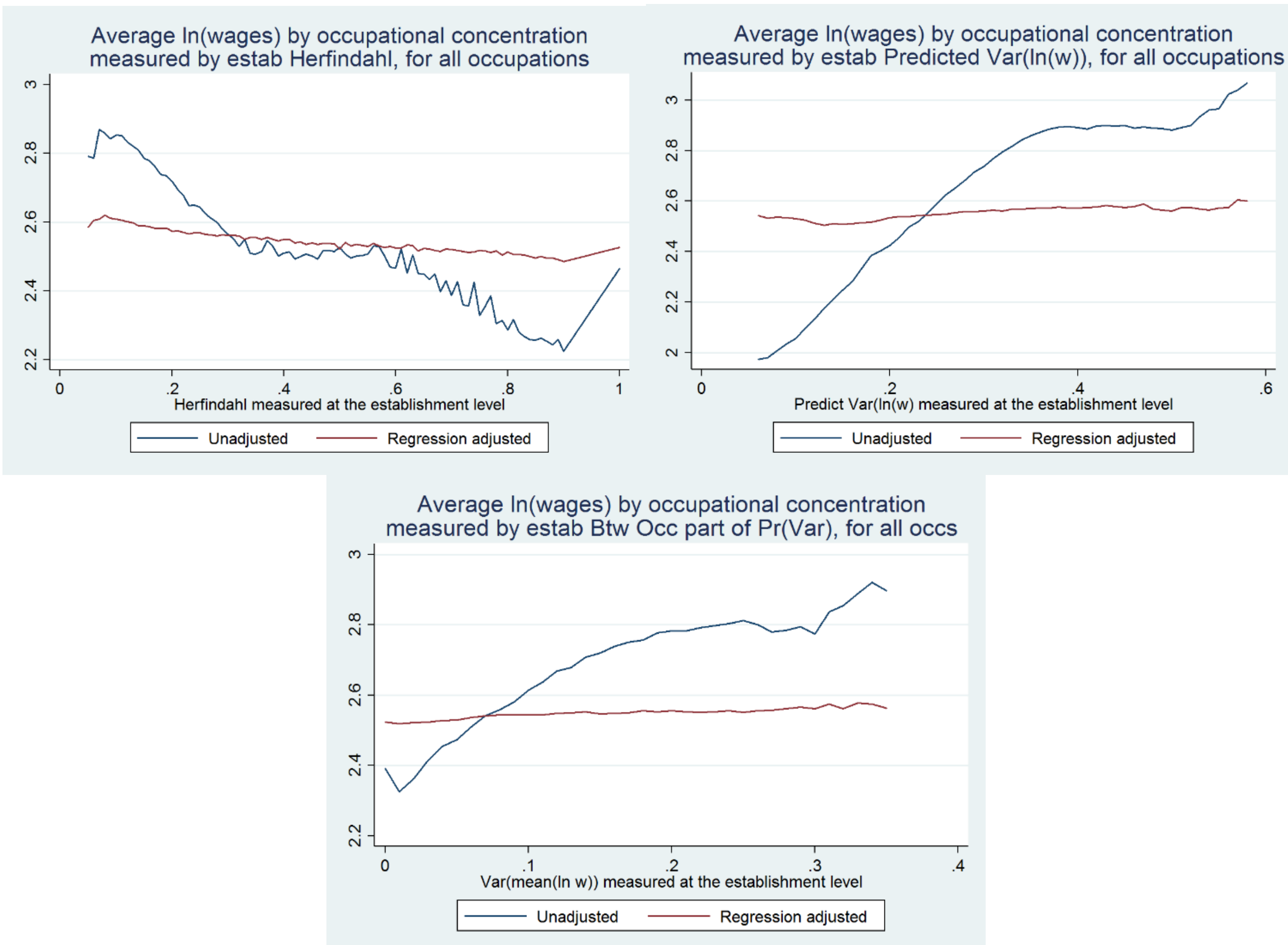
[26] Where the sample design allows it, the OES program also makes an effort to sample establishments of the same employer in the same wave of the sample. Thus, the establishments of individual large employers are likely to appear in the OES data every six panels.

for establishment characteristics and occupations are added to equation (2), or if the occupations in the lowest-paid quintile are examined separately.

Trends in occupational concentration differ between very large employers and smaller employers for smaller employers, for each measure and for all occupations as well as for the lowest-paid quintile of occupations, with and without controlling for other observable variables. By all measures, equation (4) shows small but significant increases in occupational concentration within huge employers for every measure of occupational concentration. This is the opposite of the trend found for small employers for the predicted variance of wages (and its between-establishment component) measures. [27]

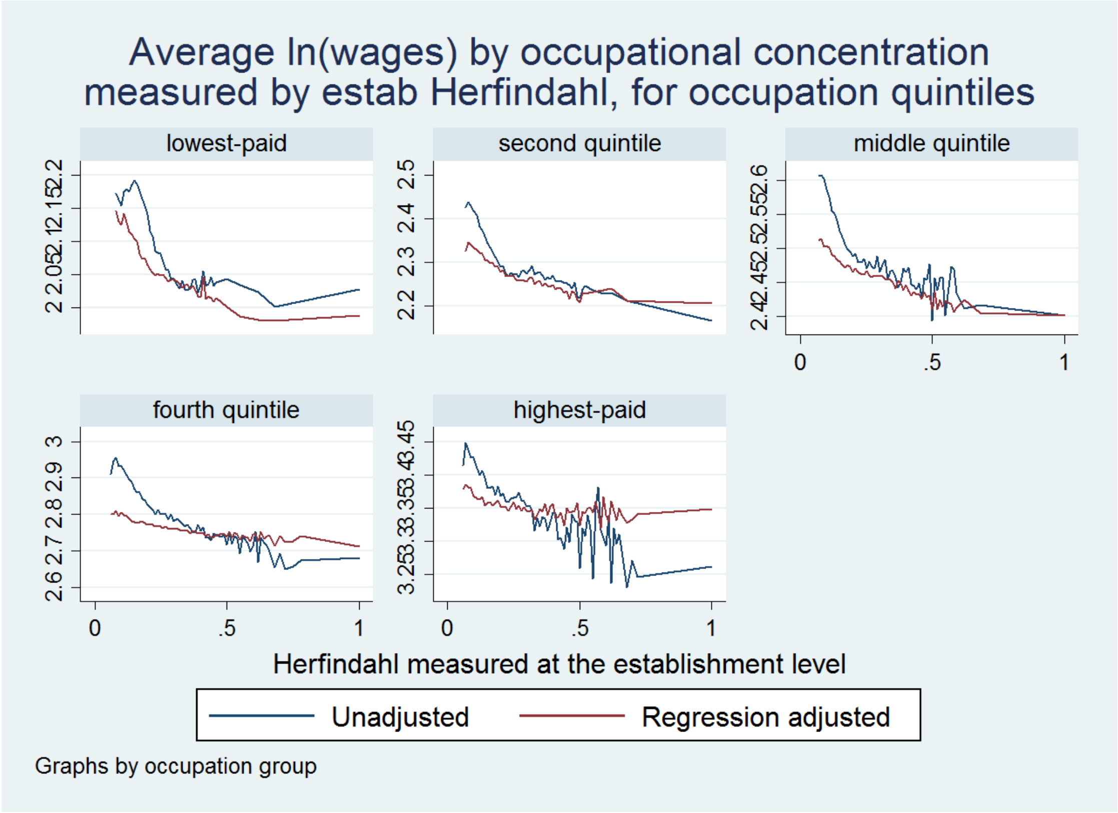
[27] Results available upon request.

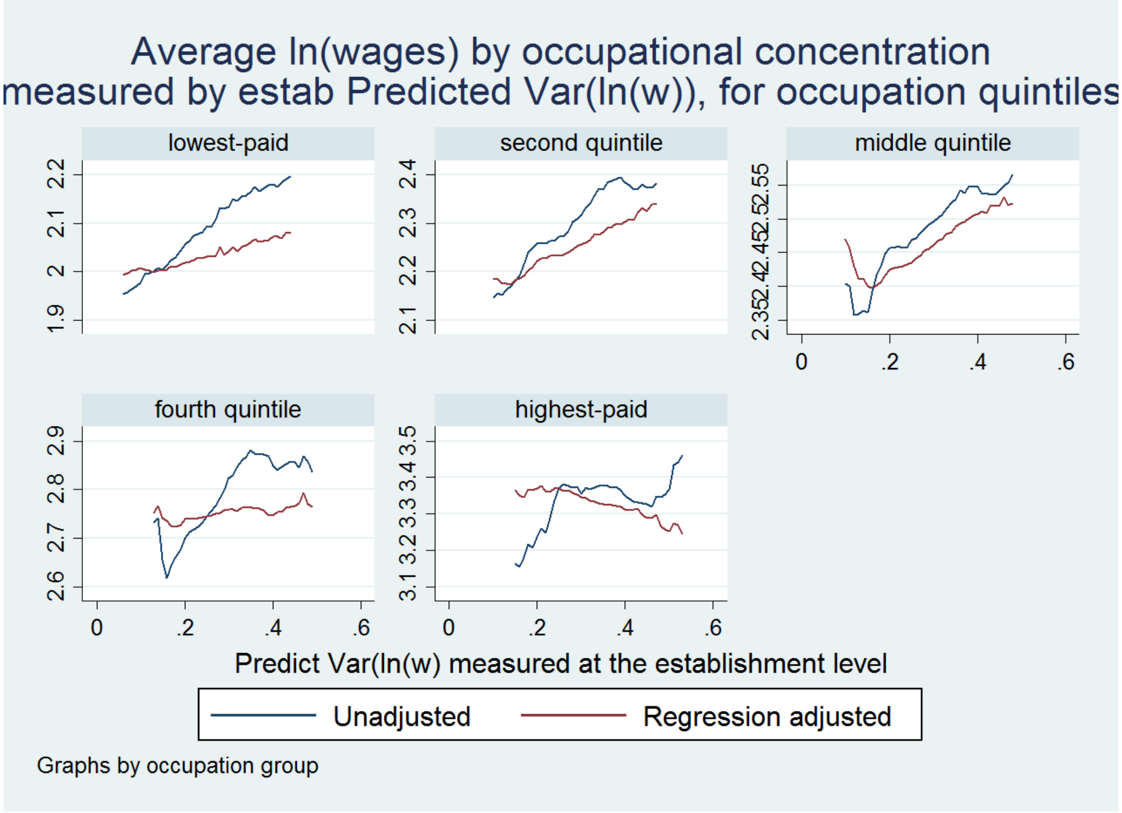
Figure 1: Relationships between Wages and Occupational Concentration

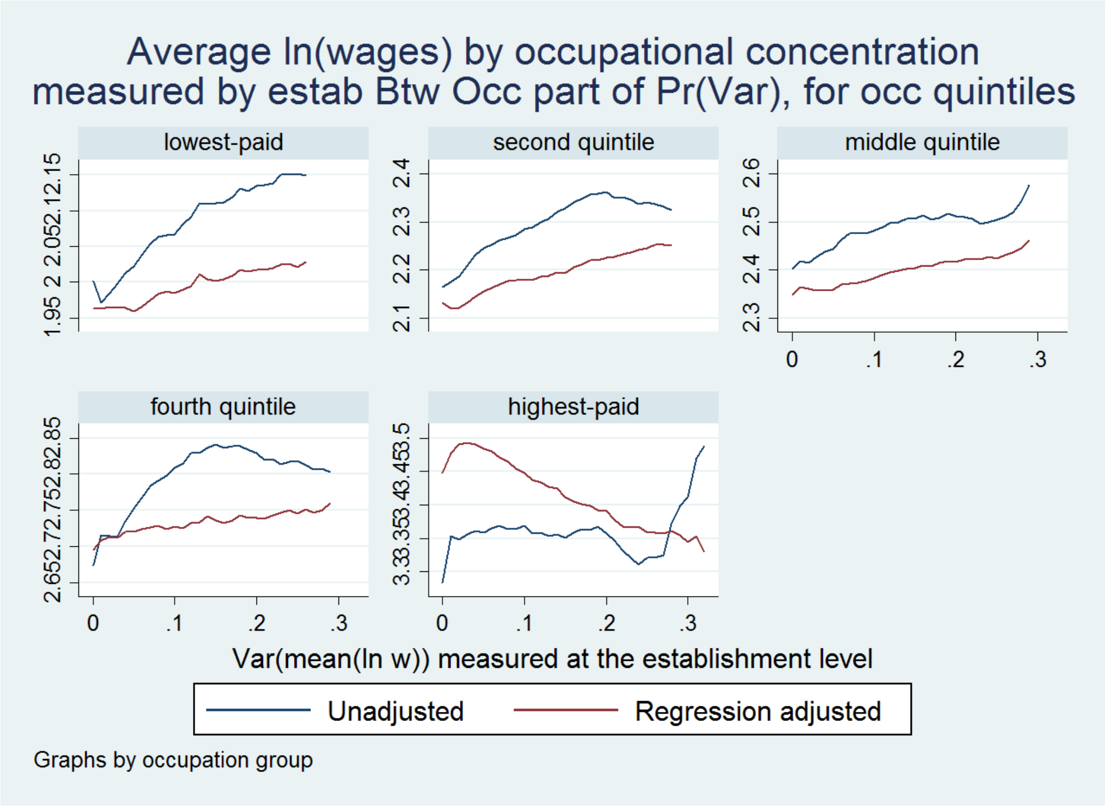


Notes: the "Average ln(wage)" coefficients plotted here are the set of α coefficients from regressions of the form $\ln(wage_{ijt}) = \sum \alpha_m I(m - .01 < Occ\ Concen_{jt} \leq m) + \delta X_{ijt} + \ \varepsilon$, where Occupation Concentration groups *m* are formed by rounding each Occupation Concentration variable to the nearest hundredth, and where X_{itj} (in "regression adjusted" estimates) includes dummy variables for each detailed occupation in the OES, 4-digit employer NAICS codes, states, and employer size classes.   Graphs are based on 46,609,394 observations at the establishment-occupation-wage interval level for November 2002 through November 2016, weighted by employment. A constant term is added to the "regression adjusted" estimates to plot them on the same scale as the "unadjusted" estimates.

Figure 2: Relationships between Wages and Occupational Concentration by occupation quintile

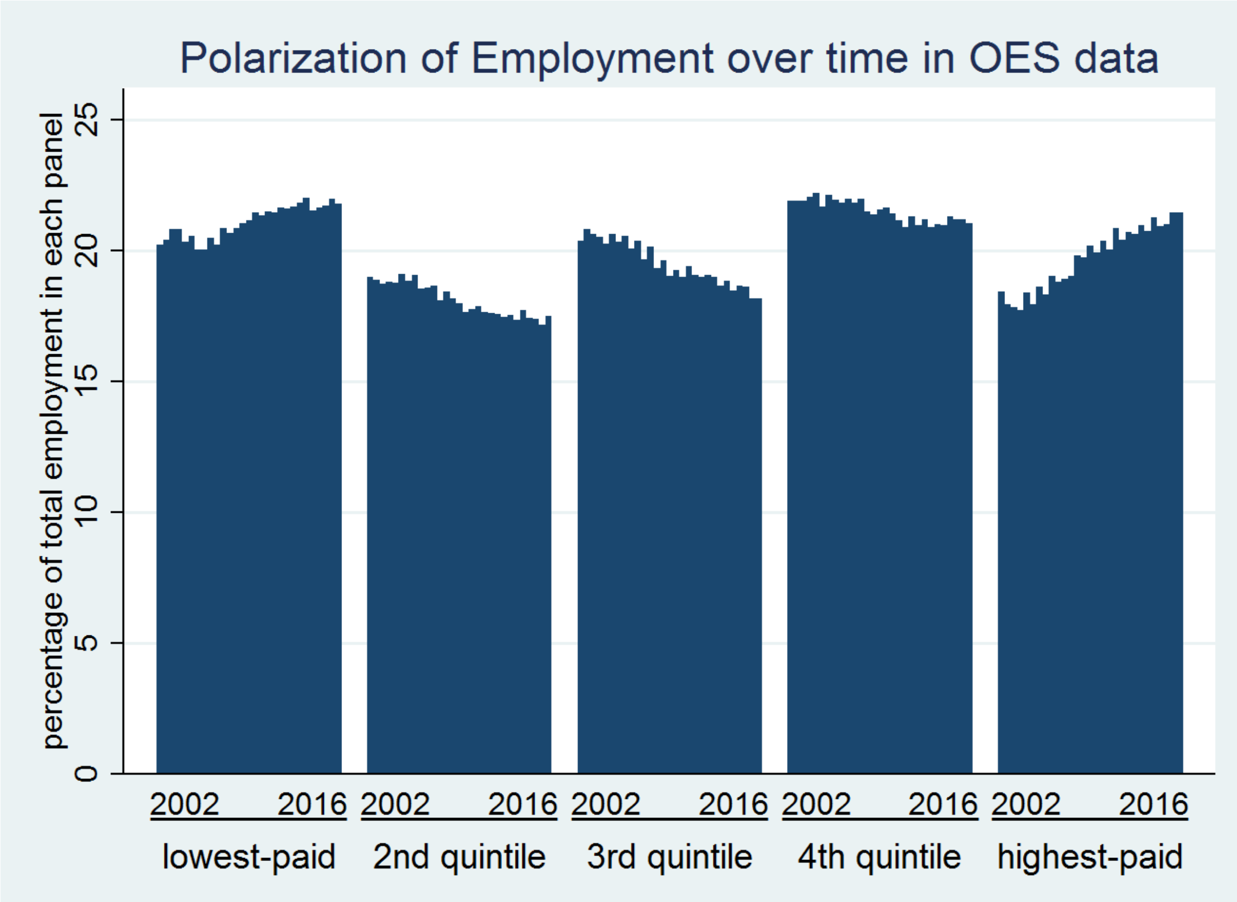






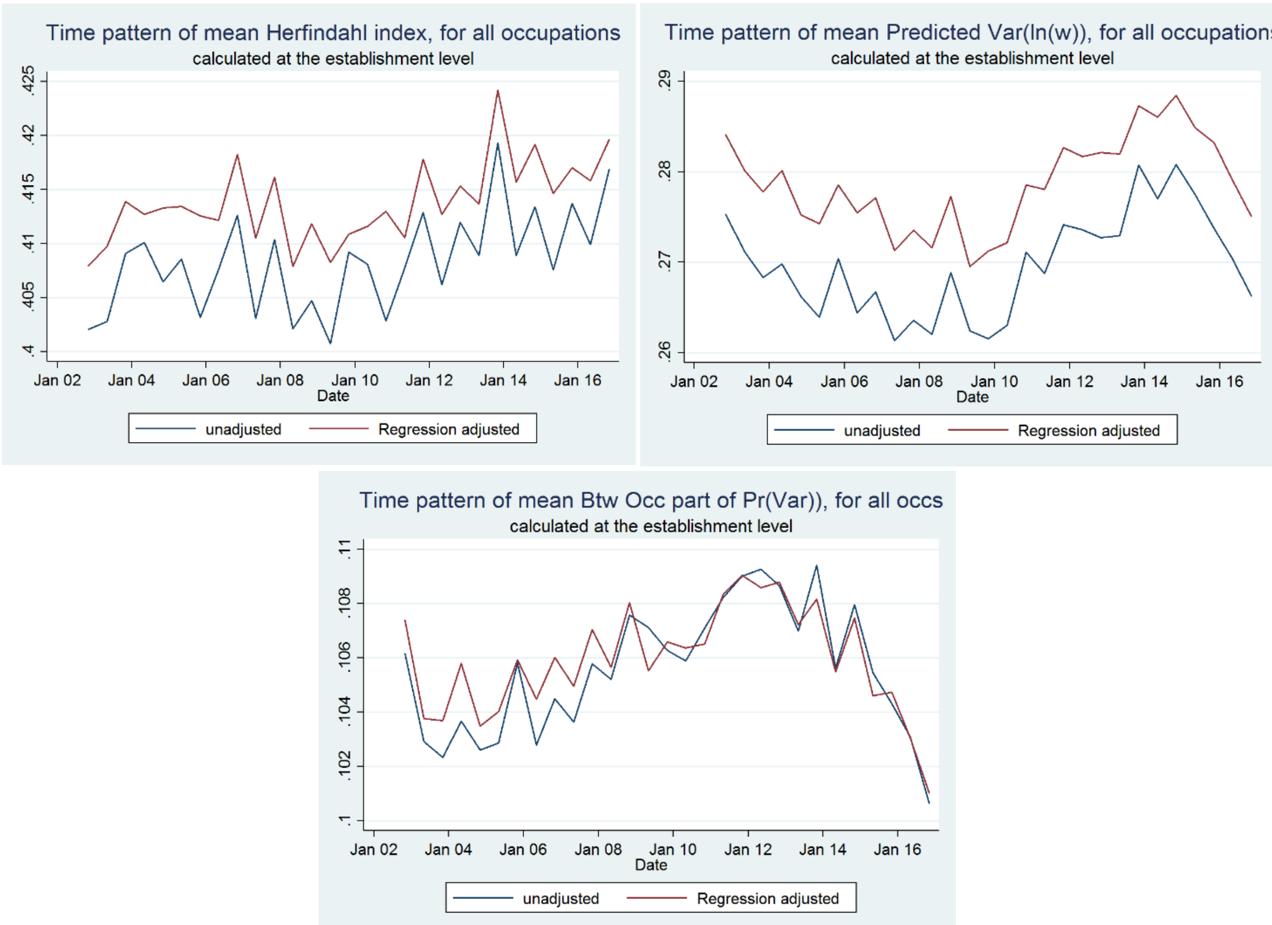
See notes for Figure 1.

Figure 3: Trends in Employment by Quintile of Occupation



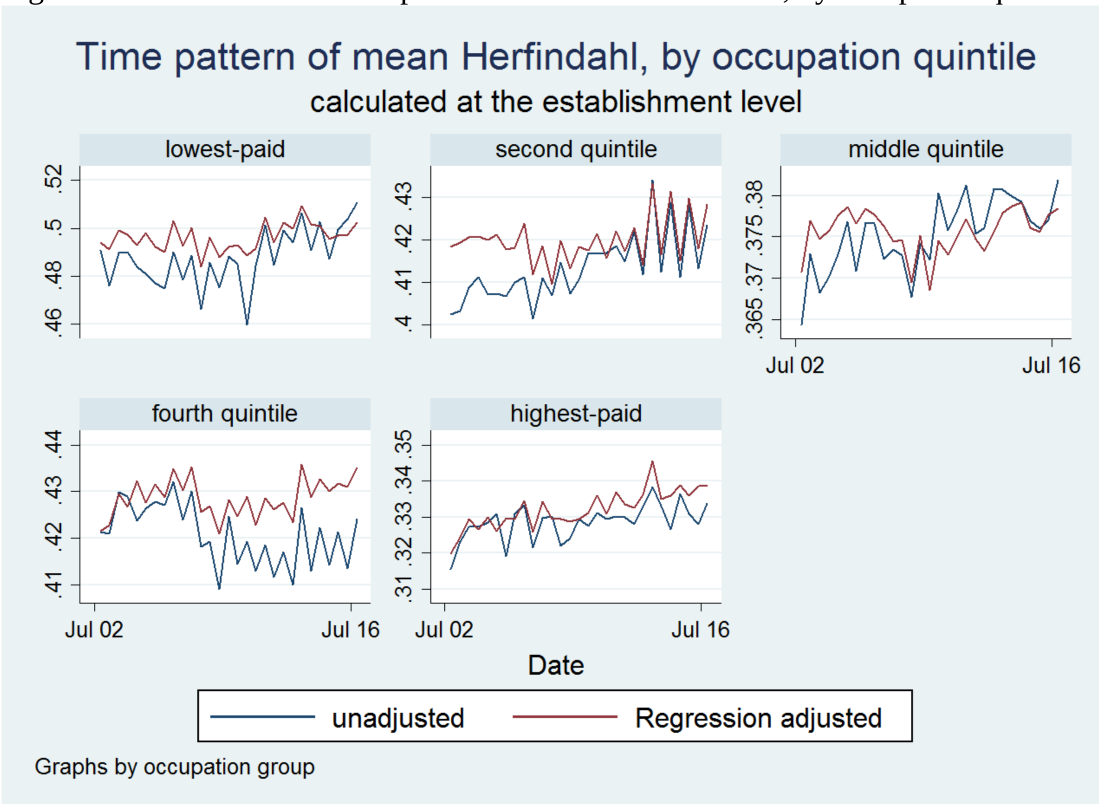
Note: The 46,609,394 observations in 29 waves of data are used to calculate overall average wage levels and employment levels (as shown in Appendix A). These are grouped into quintiles of occupation by average wage levels. This figure shows the percentage of employment in each occupational quintile in each wave of OES data, from November 2002 through November 2016.

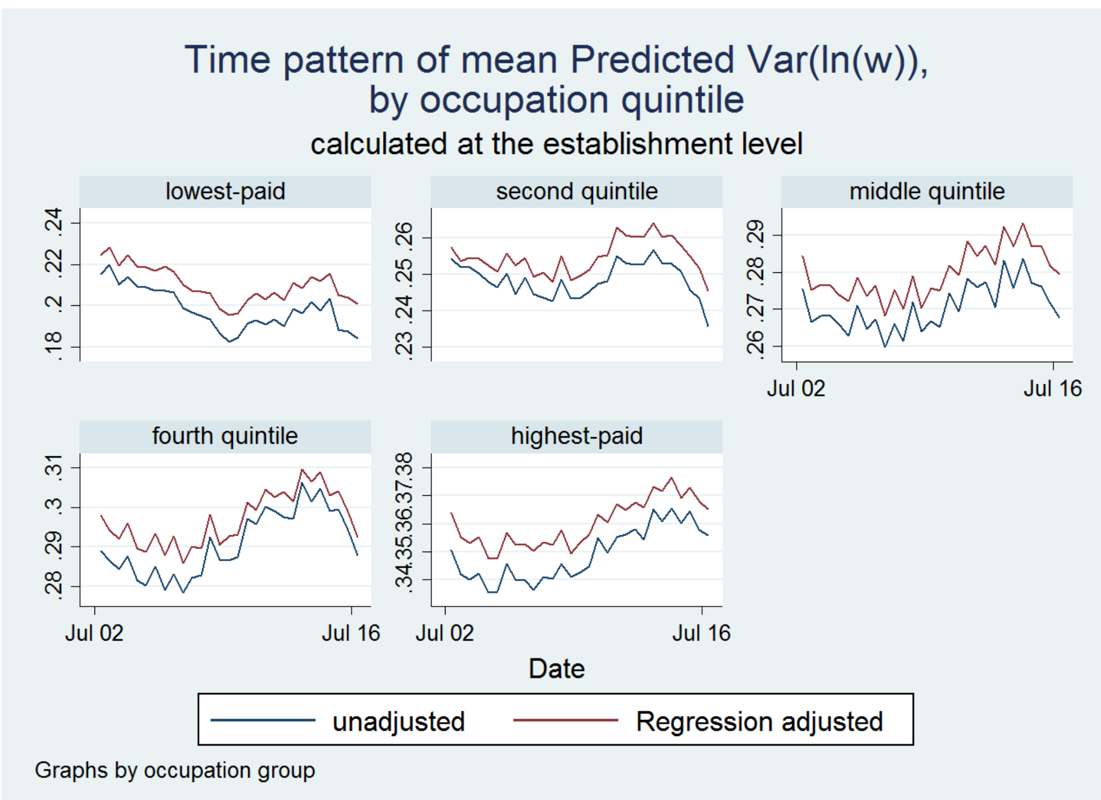
Figure 4: Trends in mean Occupational Concentration value

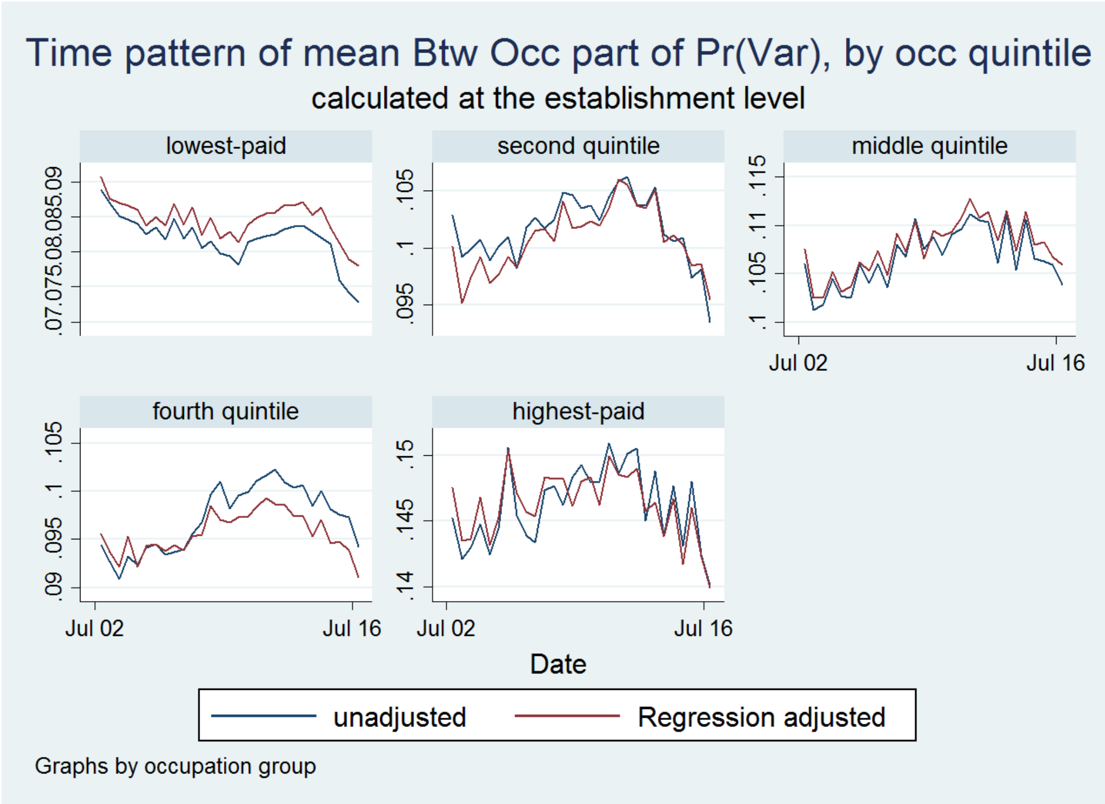


Note: These are plots of coefficients γ_t from regressions $Occupational\ Concen_{jt} = \gamma_t Survey\ Date_t + \delta X_{ijt} + \varepsilon$, where X_{itj} ("regression adjusted" estimates only) includes dummy variables for each detailed occupation in the OES, 4-digit employer NAICS codes, states, and employer size classes. Regressions are based on 46,609,394 observations at the establishment-occupation-wage interval level, weighted by employment. A constant term is added to the "regression adjusted" estimates to plot them on the same scale as the "unadjusted" estimates.

Figure 5: Trends in mean Occupational Concentration values, by occupation quintile

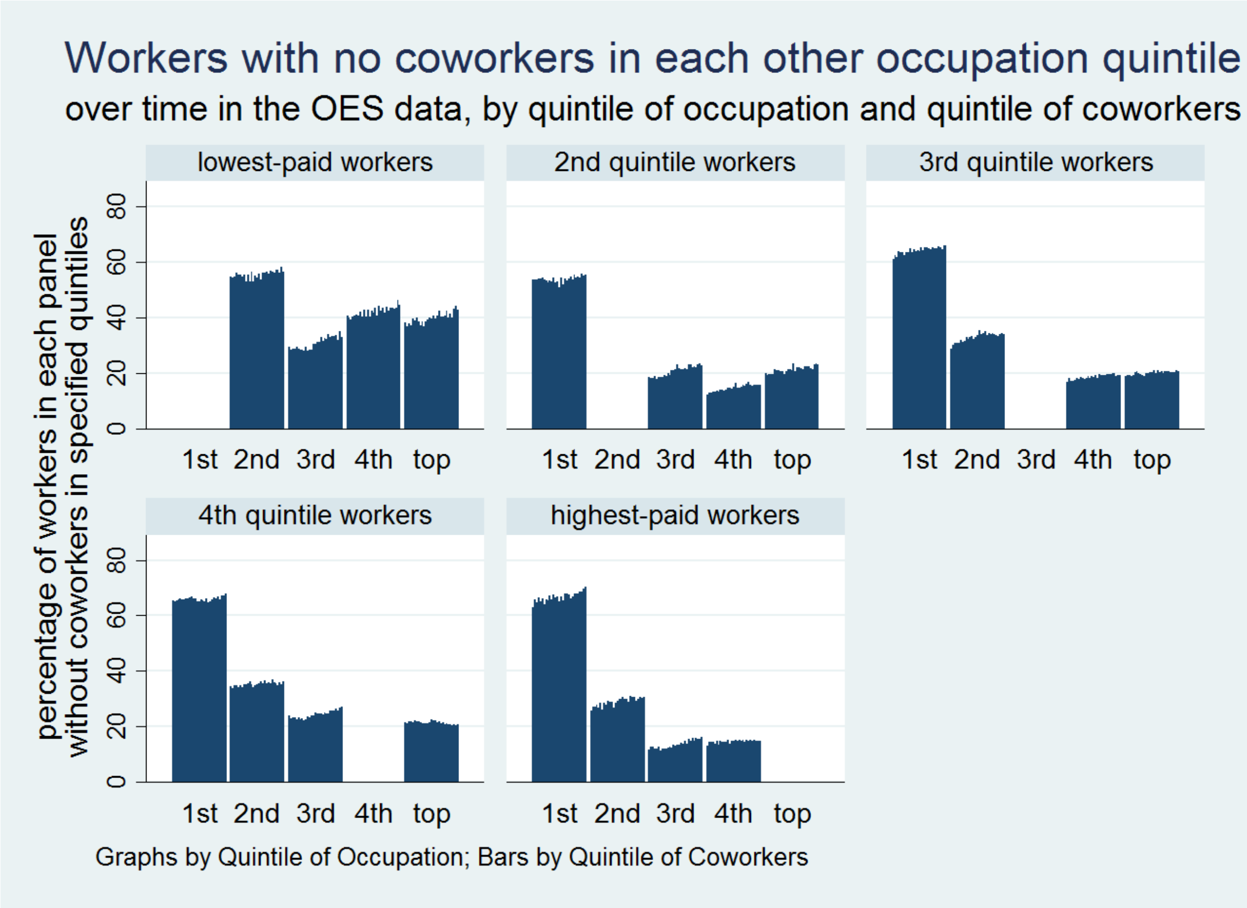






Time pattern of mean Btw Occ part of Pr(Var), by occ quintile
calculated at the establishment level
lowest-paid
second quintile
middle quintile
fourth quintile
highest-paid
.07 .075 .08 .085 .09
.095 .1 .105
.1 .105 .11 .115
.09 .095 .1 .105
.14 .145 .15
Jul 02
Jul 16
Date
unadjusted
Regression adjusted
Graphs by occupation group

Figure 6: Workers with no coworkers in other occupational quintiles over time in the OES data, by quintile of occupation and quintile of coworkers



Note: The 46,609,394 observations in 29 waves of data are used to calculate overall average wage levels and employment levels (as shown in Appendix A). These are grouped into quintiles of occupations by average wage levels. This figure shows the percentage of workers in each quintile who are employed in establishments that have no workers in each other quintile, by panel (from November 2002 to November 2016). 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 OES data.

Table 1: Summary Statistics

Variable description	Occupation by wage interval observations	Employment represented	Weighted Mean	Minimum	Maximum	Variance
OES wages, by establishment-occupation-wage interval	46,609,428	2,242,528,409	20.043	5.778	145.759	310.785
OES real wages, by estab-occupation-wage interval	46,609,428	2,242,528,409	16.203	5.209	106.443	197.447
OES real ln wage, by estab-occupation-wage interval	46,609,428	2,242,528,409	2.561	1.650	4.668	0.382
measured var(ln(wg)) of establishment	46,609,428	2,242,528,409	0.154	0.000	2.091	0.019
Herfindahl of 3-digit occupations, establishment level	46,609,428	2,242,528,409	0.408	0.032	1.000	0.063
pred var(ln(wg)) for establishment, based on occupations	46,609,394	2,242,528,409	0.270	0.018	1.016	0.012
portion of above due to variation between occupations	46,609,394	2,242,528,409	0.106	0.000	0.781	0.006
predicted mean(ln(wg)) of estab, based on occupations	46,609,394	2,242,528,409	2.572	1.807	3.851	0.113
Total employment of establishment	46,609,428	2,242,528,409	576	1	56,473	5,104,405
Date of observation	46,609,428	2,242,528,409	Nov, 2009	Nov, 2002	Nov, 2016	

Variable description	Occupation by wage interval observations	Employment represented	Fraction of employment	Establishment observations
Quintiles				
Bottom quintile of occupations	4,512,048	498,610,022	22.2%	1,298,634
Second quintile of occupations	6,742,720	404,179,707	18.0%	1,946,156
Third quintile of occupations	9,757,150	434,006,537	19.4%	2,515,660
Fourth quintile of occupations	11,881,718	486,554,202	21.7%	2,806,055
Top quintile of occupations	13,715,792	419,177,942	18.7%	2,335,591
Industries (2-digit)				
Agriculture, Forestry, Fishing and Hunting	24,742	1,209,745	0.1%	5,064
Mining, Quarrying, and Oil and Gas Extraction	278,611	10,731,557	0.5%	22,949
Utilities	362,131	11,487,328	0.5%	20,980
Construction	2,715,765	140,698,009	6.3%	320,008
Manufacturing	7,617,106	247,417,992	11.0%	404,392
Wholesale Trade	3,121,983	106,037,580	4.7%	289,829
Retail Trade	6,033,799	344,932,292	15.4%	502,403
Transportation and Warehousing	1,372,852	92,123,403	4.1%	142,832
Information	1,509,864	49,886,150	2.2%	110,529
Finance and Insurance	2,691,456	118,883,973	5.3%	227,072
Real Estate and Rental and Leasing	673,419	28,956,088	1.3%	99,687
Professional, Scientific, and Technical Services	3,514,547	143,457,542	6.4%	354,483
Management of Companies and Enterprises	1,276,951	32,276,040	1.4%	38,583
Administrative and Support and Waste Management and I	2,291,647	140,897,584	6.3%	246,780
Educational Services	1,298,489	50,396,572	2.2%	70,057
Health Care and Social Assistance	6,809,349	353,872,634	15.8%	420,820
Arts, Entertainment, and Recreation	1,062,174	37,963,004	1.7%	96,342
Accommodation and Food Services	2,078,374	237,573,956	10.6%	209,645
Other Services (except Public Administration)	1,876,169	93,726,962	4.2%	282,115
Occupations (2-digit)				
Management Occupations	5,121,383	103,271,611	4.6%	
Business and Financial Operations Occupations	3,695,454	99,137,164	4.4%	
Computer and Mathematical Occupations	1,705,143	53,350,145	2.4%	
Architecture and Engineering Occupations	1,210,073	39,732,902	1.8%	
Life, Physical, and Social Science Occupations	347,605	12,029,141	0.5%	
Community and Social Service Occupations	611,383	23,933,901	1.1%	
Legal Occupations	249,027	16,324,258	0.7%	
Education, Training, and Library Occupations	654,815	38,966,650	1.7%	
Arts, Design, Entertainment, Sports, and Media Occupatio	981,616	28,832,089	1.3%	
Healthcare Practitioners and Technical Occupations	2,195,949	134,205,129	6.0%	
Healthcare Support Occupations	732,683	75,368,512	3.4%	
Protective Service Occupations	322,367	23,877,442	1.1%	
Food Preparation and Serving Related Occupations	1,918,555	223,615,488	10.0%	
Building and Grounds Cleaning and Maintenance	1,125,051	72,733,714	3.2%	
Personal Care and Service Occupations	800,258	73,439,546	3.3%	
Sales and Related Occupations	4,545,163	297,008,580	13.2%	
Office and Administrative Support Occupations	10,417,328	375,746,893	16.8%	
Farming, Fishing, and Forestry Occupations	63,893	3,315,741	0.1%	
Construction and Extraction Occupations	1,546,401	108,571,047	4.8%	
Installation, Maintenance, and Repair Occupations	2,384,025	97,043,195	4.3%	
Production Occupations	3,467,582	174,526,262	7.8%	
Transportation and Material Moving Occupations	2,513,674	167,498,997	7.5%	

Table 2: Mean Values of Occupational Concentration for Specified Occupations and Industries, 2002-2016

		Mean Value of Occupational Concentration		
Occupation and Industry	Avg ln(wage)	Herfindahl of Occupational Concentration for the establishment	Predicted Variance of Wages for the establishment	Component due to Variation Between Occupations
Food preparation and serving (SOC 35)				
within Food Services (NAICS 722) – 79%	1.99	0.496	0.138	0.058
within all other industries – 21%	2.11	0.252	0.245	0.125
Janitors (SOC 372011)				
within Janitorial Services (NAICS 561720) – 46%	2.04	0.842	0.141	0.043
within all other industries – 54%	2.16	0.329	0.267	0.121
Security Guards (SOC 339032)				
within Security Guard Srvcs (NAICS 561612) – 62%	2.16	0.883	0.158	0.030
within all other industries – 38%	2.32	0.322	0.267	0.123
Truck Drivers (SOC 53303)				
within Truck Transportation (NAICS 484) – 30%	2.69	0.636	0.204	0.040
within all other industries – 70%	2.45	0.379	0.248	0.085
Accountants (SOC 132011)				
within Accounting Services (NAICS 541211) – 25%	3.22	0.574	0.307	0.085
within all other industries – 75%	3.16	0.285	0.343	0.138
Computer Occupations (SOC 151)				
within Computer Services (NAICS 5415) – 27%	3.34	0.588	0.272	0.059
within all other industries – 73%	3.31	0.302	0.329	0.122
Engineers (SOC 172)				
within Engineering Services (NAICS 54133) – 22%	3.40	0.401	0.264	0.097
within all other industries – 78%	3.44	0.249	0.309	0.133
Lawyers (SOC 231011)				
within Law Offices (NAICS 54111) – 84%	3.76	0.411	0.544	0.284
within all other industries – 16%	3.87	0.277	0.409	0.163

Table 3: Regressions of log wages on measures of Occupational Concentration

All unimputed OES private-sector data from November 2002-November 2016

Occupational Concentration Variable	Herfindahl of occupational concentration for the establishment		Predicted Variance of Wages for the establishment		Component due to Variation Between Occupations	
With survey-date fixed effects						
Coefficient on OccConcen	-0.496	***	2.285	***	1.779	***
(standard error)	-0.001		(0.002)		(0.002)	
Coefficient on OccConcen * Date	-0.050	***	0.156	***	0.146	***
(standard error)	-0.001		(0.002)		(0.003)	
With survey-date and 6-digit occupation fixed effects						
Coefficient on OccConcen	-0.213	***	0.524	***	0.461	***
(standard error)	0.000		(0.001)		(0.003)	
Coefficient on OccConcen * Date	0.003	***	-0.048	***	-0.025	***
(standard error)	-0.001		(0.001)		(0.002)	
With survey-date, 6-digit occupation, 5-digit NAICS, & state fixed effects, and estab size						
Coefficient on OccConcen	-0.107	***	0.234	***	0.165	***
(standard error)	0.000		(0.001)		(0.001)	
Coefficient on OccConcen * Date	0.008	***	-0.055	***	-0.036	***
(standard error)	0.000		(0.001)		(0.002)	
N	46,609,394		46,609,394		46,609,394	

Notes: These regressions are of the form $Ln(wage_{ijt}) = \alpha OccConcen_{jt} + \beta OccConcen_{jt} * Date + \delta X_{ijt} + \varepsilon$, where Date is measured in decades since November 1, 2002, X includes survey 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 a continuous measure of establishment size). "*" indicates p<0.05; "**" indicates p<0.01; and "***" indicates p<0.001. Regressions are based on 46,609,394 observations at the establishment-occupation-wage interval level, weighted by employment.

Table 4: Regressions of log wages on measures of Occupational Concentration by Occupations grouped into Quintiles of Employment

Wage regressions for quinitles of unimputed OES private-sector data from Nov 2002 - Nov 2016

Occupations are ranked by average wage and classified into quintiles as shown in Appendix A

Occupational Concentration Variable	Herfindahl of occupational concentration for the establishment		Predicted Variance of Wages for the establishment		Component due to Variation Between Occupations	
Occupations in the lowest quintile of average wages (4,512,045 observations)						
Coefficient with only date fixed effects	-0.229	***	0.882	***	0.828	***
(standard error)	(0.001)		(0.003)		(0.004)	
Coefficient with full set of controls	-0.145	***	0.358	***	0.364	***
(standard error)	(0.001)		(0.003)		(0.004)	
Occupations in the second quintile of average wages (6,742,713 observations)						
Coefficient with only date fixed effects	-0.254	***	0.722	***	0.730	***
(standard error)	(0.001)		(0.003)		(0.003)	
Coefficient with full set of controls	-0.159	***	0.377	***	0.433	***
(standard error)	(0.001)		(0.002)		(0.003)	
Occupations in the middle quintile of average wages (9,757,141 observations)						
Coefficient with only date fixed effects	-0.199	***	0.579	***	0.546	***
(standard error)	(0.001)		(0.002)		(0.003)	
Coefficient with full set of controls	-0.122	***	0.295	***	0.323	***
(standard error)	(0.001)		(0.002)		(0.003)	
Occupations in the fourth quintile of average wages (11,881,704 observations)						
Coefficient with only date fixed effects	-0.249	***	0.647	***	0.607	***
(standard error)	(0.001)		(0.003)		(0.004)	
Coefficient with full set of controls	-0.074	***	0.096	***	0.131	***
(standard error)	(0.001)		(0.003)		(0.003)	
Occupations in the fifth quintile of average wages (13,715,791 observations)						
Coefficient with only date fixed effects	-0.206	***	0.463	***	0.306	***
(standard error)	(0.001)		(0.003)		(0.004)	
Coefficient with full set of controls	0.009	***	-0.153	***	-0.354	***
(standard error)	(0.001)		(0.003)		(0.003)	

Notes: These regressions are of the form $Ln(wage_{ijt}) = \alpha OccConcen_{jt} + \beta OccConcen_{jt} * Date + \delta X_{ijt} + \varepsilon$, where Date is measured in decades since November 1, 2002, X includes survey 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 a continuous measure of establishment size). "*" indicates p<0.05; "**" indicates p<0.01; and "***" indicates p<0.001. Regressions are based on 46,609,394 observations at the establishment-occupation-wage interval level, weighted by employment

Table 5: Changes in mean values of Occupational Concentration over time

All unimputed OES private-sector data from November 2002-November 2016

Occupational Concentration Variable	Herfindahl of occupational concentration for the establishment		Predicted Variance of Wages for the establishment		Component due to Variation Between Occupations	
Raw time trends in occupational concentration						
Coefficient (per decade)	0.0057	***	0.0059	***	0.0017	***
(standard error)	(0.00009)		(0.00004)		(0.00003)	
Time trends, controlling for 6-digit occupation						
Coefficient (per decade)	0.0061	***	0.0036	***	-0.0003	***
(standard error)	(0.00008)		(0.00003)		(0.00002)	
Time trends, controlling for 6-digit occupation, 5-digit NAICS codes, size class, size, & state						
Coefficient (per decade)	0.0049	***	0.0048	***	0.0004	***
(standard error)	(0.00006)		(0.00003)		(0.00002)	
N	46,609,394		46,609,394		46,609,394	

Note: These are coefficients α from regressions of the form $Occupational\ Concen_{jt} = \gamma Survey\ Date_t + \theta I(May\ Survey) + \delta X_{ijt} + \varepsilon$, where X_{itj} 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. "*" indicates p<0.05; "**" indicates p<0.01; and "***" indicates p<0.001. Regressions are based on 46,609,394 observations at the establishment-occupation-wage interval level, weighted by employment.

Table 6: Changes in mean values of Occupational Concentration over time by Occupations grouped into Quintiles of Employment

Occupational Concentration Variable	Herfindahl of occupational concentration for the establishment		Predicted Variance of Wages for the establishment		Component due to Variation Between Occupations	
Occupations in the lowest quintile of average wages (4,512,045 observations)						
Coefficient with no controls (per decade)	0.015	***	-0.017	***	-0.005	***
(standard error)	(0.00030)		(0.00011)		(0.00008)	
Coefficient with controls (per decade)	0.005	***	-0.012	***	-0.003	***
(standard error)	(0.00020)		(0.00008)		(0.00007)	
Occupations in the second quintile of average wages (6,742,713 observations)						
Coefficient with no controls (per decade)	0.013	***	-0.001	***	-0.0002	**
(standard error)	(0.00024)		(0.00009)		(0.00007)	
Coefficient with controls (per decade)	0.004	***	0.003	***	0.002	***
(standard error)	(0.00016)		(0.00007)		(0.00005)	
Occupations in the middle quintile of average wages (9,757,141 observations)						
Coefficient with no controls (per decade)	0.007	***	0.008	***	0.004	***
(standard error)	(0.00017)		(0.00008)		(0.00006)	
Coefficient with controls (per decade)	0.001	***	0.010	***	0.004	***
(standard error)	(0.00013)		(0.00006)		(0.00005)	
Occupations in the fourth quintile of average wages (11,881,704 observations)						
Coefficient with no controls (per decade)	-0.008	***	0.014	***	0.005	***
(standard error)	(0.00018)		(0.00006)		(0.00005)	
Coefficient with controls (per decade)	0.003	***	0.010	***	0.001	***
(standard error)	(0.00013)		(0.00004)		(0.00004)	
Occupations in the fifth quintile of average wages (13,715,791 observations)						
Coefficient with no controls (per decade)	0.007	***	0.019	***	0.001	***
(standard error)	(0.00014)		(0.00006)		(0.00005)	
Coefficient with controls (per decade)	0.010	***	0.016	***	-0.002	***
(standard error)	(0.00011)		(0.00004)		(0.00004)	

Note: These are coefficients α from regressions of the form $Occupational\ Concen_{jt} = \gamma Survey\ Date_t + \theta I(May\ Survey) + \delta X_{ijt} + \varepsilon$, where X_{itj} 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. "*" indicates p<0.05; "**" indicates p<0.01; and "***" indicates p<0.001. Regressions are based on 46,609,394 observations at the establishment-occupation-wage interval level, weighted by employment.

Table 7: Explanatory Power of Selected Characteristics on Establishment-level ln Wage Inequality: November 2002, November 2009, and November 2016

	R^2 values from regressions of establishment-level mean wages		
Explanatory Variables	Nov 2002	Nov 2009	Nov 2016
(1) 5-digit NAICS, size class, state fixed effects, and continuous size	.700	.666	.671
(2) Herfindahl of occupational concentration	.117	.107	.127
(3) Predicted Variance of ln(wages) for the establishment based on its occupational composition	.244	.264	.351
(4) Component of (3) due only to variation in mean wages between the occupations employed in the establishment	.088	.093	.124
(5) Predicted Mean(ln(wages)) for the establishment based on the occupations employed at the establishment	.697	.699	.731
(6) (1) + (2)	.713	.675	.681
(7) (1) + (3)	.733	.700	.715
(8) (1) + (4)	.712	.674	.680
(9) (1) + (5)	.819	.811	.826
(10) (1) + (2) + (3) +(4)	.749	.723	.750
(11) (1) + (2) + (3) + (4) +(5)	.820	.812	.827

Notes: R^2 values from establishment-level regressions of the form mean $\ln(wage)_j = \lambda X_j$, where X variables are as listed in the first column, and observations are weighted by the employment in each establishment. Regressions are based on observations from 137,333 establishments in November 2002, 133,912 establishments in November 2009, and 124,366 establishments in November 2016.

Table 8: Estimates of November 2015 Wage Variance, Reweighted to Characteristics in Fall 1999, Selected combinations of characteristics

		Overall		Between Estabs		Within Estabs		Hourly Wage at selected percentiles			Percentile ranges		
	Variables used in Dinardo-Fortin-Lemiuex reweighting exercise	Variance	Variance Growth Explained	Variance	Variance Growth Explained	Variance	Variance Growth Explained	10th	50th	90th	50-10	90-50	90-10
(1)	No reweighting--actual May & Nov 2016 wage distribution	0.3861		0.2325		0.1536		$6.55	$12.24	$32.62	$5.69	$20.38	$26.07
(2)	4-digit industry	0.3788	42%	0.2261	31%	0.1527	-26%	$6.55	$12.23	$32.06	$5.68	$19.83	$25.51
(3)	50-state location	0.3879	-11%	0.2338	-6%	0.1541	17%	$6.57	$12.26	$32.79	$5.69	$20.53	$26.22
(4)	Establishment size category	0.3843	11%	0.2314	5%	0.1528	-22%	$6.53	$12.17	$32.38	$5.64	$20.21	$25.85
(5)	3-digit occupation	0.3681	104%	0.2192	65%	0.1489	-138%	$6.55	$12.06	$31.23	$5.51	$19.17	$24.68
(6)	occupation quintile	0.3650	123%	0.2179	71%	0.1471	-193%	$6.55	$12.03	$31.03	$5.48	$19.00	$24.48
(7)	decile of Herfindahl index * quintile	0.3644	126%	0.2165	78%	0.1479	-169%	$6.57	$12.05	$31.05	$5.48	$19.00	$24.48
(8)	decile of predicted variance of ln(wages) * quintile	0.3613	144%	0.2134	93%	0.1479	-167%	$6.56	$12.06	$31.06	$5.50	$19.00	$24.50
(9)	decile of between-occupations portion of predicted variance * quintile	0.3643	127%	0.2163	79%	0.1480	-165%	$6.61	$12.09	$31.07	$5.48	$18.98	$24.46
(10)	Combination of (2)+(3)+(4)+(5)+(7)+(8)	0.3714	85%	0.2163	79%	0.1551	45%	$6.58	$12.12	$31.39	$5.54	$19.27	$24.81
(11)	Most between-estab variance growth explained: Combination of (2)+(7)+(8)	0.3635	131%	0.2131	95%	0.1505	-92%	$6.56	$12.10	$31.11	$5.54	$19.01	$24.55
(12)	Most variance and most between-estab variance explained without (7), (8), or (9): Combination of (4)+(6)	0.3647	124%	0.2175	73%	0.1472	-190%	$6.54	$12.00	$30.94	$5.46	$18.94	$24.40
(13)	No (or only) Quintile 5 employment in establishment * quintile	0.3642	127%	0.2162	79%	0.1480	-166%	$6.57	$12.04	$31.03	$5.47	$18.99	$24.46
(14)	Even more btw-estab var growth explained: Combination of (2)+(8)+(13)	0.3627	136%	0.2121	99%	0.1506	-87%	$6.57	$12.11	$31.13	$5.54	$19.02	$24.56

See Section IVb for an explanation. Reweighting follows Dinardo, Fortin and Lemieux 1996.