

Outsourcing, Occupational and Industrial Concentration*

Nicholas Bloom (Stanford), Audrey Guo (Stanford) and Brian Lucking (Stanford)

November 2018

Abstract

We use data from the U.S. Bureau of Labor Statistics and U.S. Census Bureau to investigate the concentration of firms, occupations, and industries over time. We find strong evidence for increased concentration of firms in terms of a secular decrease in the number of occupations and industries in which firms are active. The mean number of 5-digit occupations per establishment has fallen from 6.5 to 5.5 since 2000, with the top 3 occupations now accounting for over 85% of total establishment employment. Firm employment and payroll is increasingly concentrated in a few core industries. We argue that the rise of outsourcing and pressure on firms to focus on their core competencies is driving this concentration of activities within firms. Finally, we argue this could play a role for the rising segregation of employees by income and education across firms.

JEL No. L2, M2, O32, O33.

Keywords: occupational concentration, outsourcing, core competencies

*Acknowledgments: We would like to thank National Science Foundation and the Alfred Sloan Foundation for financial support.

1 Introduction

Over the last thirty years U.S. business schools and the business media have been increasingly focused on outsourcing and core competencies. Outsourcing is argued to allow firms to spin off noncritical activities like security, catering, cleaning, accounting, or real-estate to external providers that can potentially perform these tasks more cheaply or more effectively, and has been gaining interest since the early 1990s.¹ The related concept of core competencies refers to the idea that firms have a set of activities they can perform well, and outside these “core competencies” their ability starts to deteriorate (Prahalad and Hamel (1990)). This is either because of a lack of innate ability - for example, the claim that Walmart should focus on retailing, Marriot on hotels, and Starbucks on coffee shops - or potentially due to managerial distraction when firms span multiple industries. More recently a number of papers have linked this phenomena to growing income inequality, arguing that firms can cut wages for peripheral activities by outsourcing them (see e.g. Song et. al (2019); Goldschmidt and Schmieder (2017)).

This paper investigates the extent of outsourcing by using two types of data. The first looks at occupational concentration within firms using the Bureau of Labor Statistics (BLS) Survey of Occupational Employment (OES). The OES is unique in the U.S. in having detailed panel data on the occupational breakdown of up to 200 occupations per firm in a twice-yearly panel covering around 200,000 establishments a year. Using this data we show a gradual increase in occupational concentration since the current occupational classification system began in 1999. The second uses U.S. Census data on firm industry mix based on two measures - firstly, the employment shares of firms across industries using the primary industry of their establishments from the Longitudinal Business Database (LBD), and secondly for manufacturing firms the concentration of output by industry using detailed product information from the quinquennial (once every five years) Census of Manufacturing (CMF). Both datasets show a similar increase in industrial concentration within firms since the 1970s. We also confirm that in a panel of Compustat data, which covers U.S. publicly listed firms.

In section 2 we discuss our BLS and Census data, in section 3 we discuss our results on rising occupational concentration and in section 4 our results on industrial concentration.

2 Data

2.1 Bureau of Labor Statistics Occupational Employment Survey

The Occupational Employment Survey is a semi-annual survey of U.S. business establishments that collects data on the distribution and earnings of occupations within establishments. To reduce respondent burden, each establishment is surveyed at most once every 3 years, and the BLS’ published occupation statistics are calculated from 3-year waves. Employers are asked to categorize each of their

¹As one measure of this the frequency of the word “Outsourcing” in books according to Google Ngrams starts to gradually rise from zero in the mid-1970s and then accelerates upwards beginning in the late 1980s)

workers into 800+ detailed occupations based on the Standard Occupational Classification (SOC) system, and to further assign each worker into one of 12 wage intervals that are updated periodically to keep up with inflation. In 2006 for example, the hourly wage intervals spanned from under \$7.50, \$7.50-\$9.49, \$9.50-\$11.99, ... up to \$80 or more. The same intervals are simply multiplied by 2,080 (the number of hours in a typical work year) for salaried workers, but the employer makes no designation as to whether the assignment is based on an hourly or salary wage. Wages include base rate pay, cost-of-living allowances, guaranteed pay, hazardous-duty pay, incentive pay such as commissions and production bonuses, tips, and on-call pay. They do not however include employer costs of supplementary benefits or overtime pay.

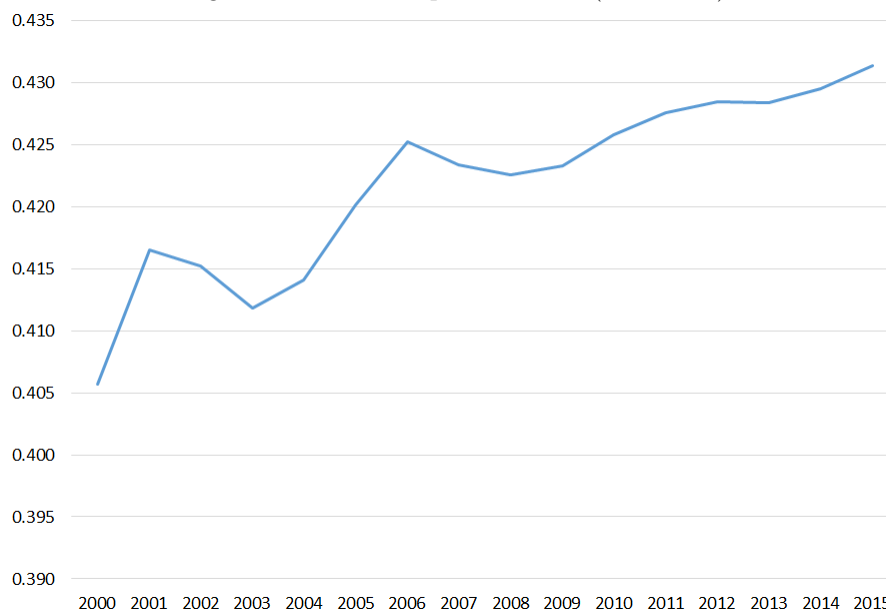
Employees are assigned to occupations based on the work they perform and not on the basis of their education or training. For workers who perform duties of multiple occupations, employers are asked to report the one that requires the highest level of skill. In this project we aggregate detailed occupations to the 5-digit SOC level. For example, “Packers and Packagers, Hand” would be aggregated with three other occupations to form “Laborers and Material Movers, Hand”. For some occupations, such as Customer Service Representatives or Food Preparation Workers, the detailed occupation code is synonymous with the 5-digit SOC group. Out of the establishments sampled, response rates range from 75 to 85%. Because data are imputed for each nonresponding establishment, we drop imputed establishments from our sample. Furthermore, we drop any establishments smaller than 5 employees to remove units too small to generate meaningful measures of concentration.

2.2 Census Data

Our measures of the industrial concentration of firms are constructed using data from the Census’s Longitudinal Business Database (LBD). The LBD covers the universe of employer business establishments and contains annual data on the industry classification and parent firm of each establishment, as well each establishment’s payroll and employment in the payroll period containing the week of March 12. The data are available for the years 1976 through 2015. To avoid outliers we drop any establishment whose mean wage is either greater than \$250,000 in 2015 dollars, or less than the full-time federal minimum wage.

In order to study the industrial concentration of *firms*, we aggregate the establishment-level LBD to the firm-level. The first step in this process is to assign each of the firm’s establishments to a unique industry. We define an industry as a 6-digit NAICS code using the longitudinally-consistent definition from Fort and Klimek (2018). We then calculate the share of the firm’s total employment (or total payroll) in each industry. The measures of industrial concentration we consider are the share of firm employment in the firm’s largest industry, the employment share in the firm’s 3 largest industries, and the employment HHI over all of the firm’s industries. We also construct analogous measures for payroll shares. Finally, we collapse the data to create a firm-year panel dataset containing all firms in the U.S. from 1976 through 2015, excluding outliers as described in the preceding paragraph.

Figure 1: Mean Occupational HHI (1999-2016)

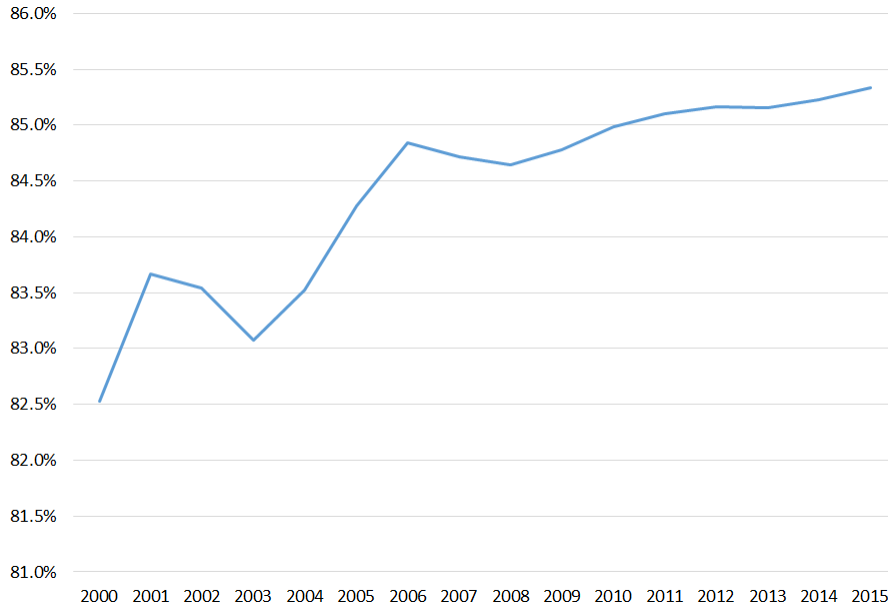


Notes: Occupational HHI is calculated for each establishment, with occupations defined at SOC-5 level. Unweighted means are calculated as 3-year averages.

3 Occupational Concentration

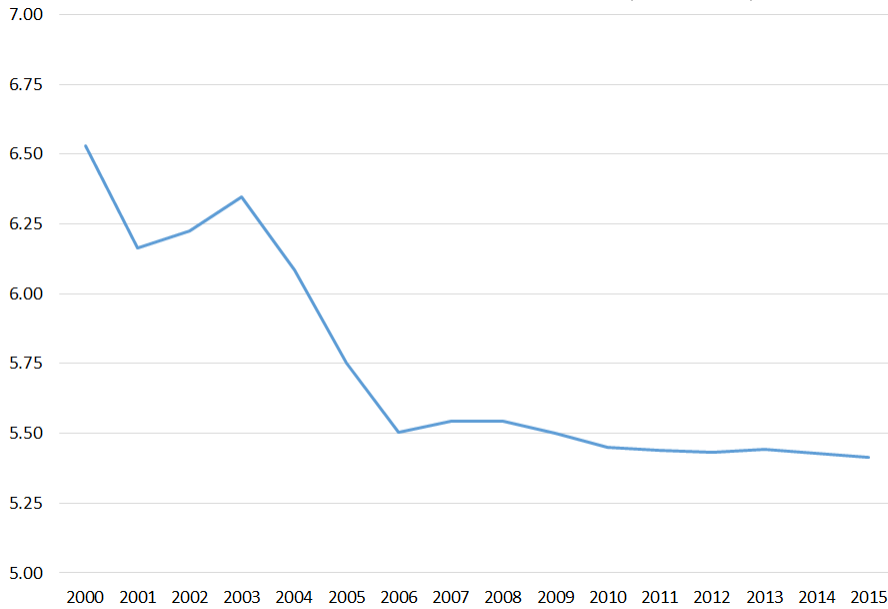
Using OES microdata from 1999 to 2016, we construct a number of different occupational concentration measures and study their change over time. We aggregate occupations into 5-digit SOC groups, and then calculated the following concentration measures at the establishment level: occupational HHI, employment share of each establishment's top 3 occupations, and number of occupations per establishment. Next, we take 3-year averages of each of these concentration measures, weighting by survey weight. This is in line with BLS' methodology of using 3-year panels to create published estimates, as establishments are only surveyed at most once every 3 years (to reduce respondent burden). Figures 1 through 3 plot the averages of each measure over time.

Figure 2: Mean Employment Share in Top 3 Occupations (1999-2016)



Notes: Top 3 Share calculated by summing the employment shares of the top 3 SOC-5 occupations in each establishment. Unweighted means are calculated as 3-year averages.

Figure 3: Mean Number of Occupations (1999-2016)



Notes: Number of occupations are calculated as the number of unique SOC-5 occupations in each establishment. Unweighted means are calculated as 3-year averages.

3.1 Regression Analysis

We also investigate this relationship through regression analyses that allow us to control for additional factors such as size and industry. Table 1 estimates a year trend on log occupational HHI over the sample period from 1999 to 2016. Column 1 estimates the baseline relationship over the entire analysis sample, and suggests that occupational HHI increased by 7.5% over the period from 1999 to 2016. This estimate is robust to excluding government agencies and small specialized NAICS sectors. These excluded sectors are Agriculture (11), Mining (21), Real Estate (53), Education (61), Other Services (81), and Public Administration (92). Further controlling for log employment, NAICS industry fixed effects, and establishment fixed effects decreases the estimated magnitude, but the year trend remains positive and statistically significant.

Table 1: Occupational HHI (1999 – 2016)

	(1)Log(HHI)	(2)Log(HHI)	(3)Log(HHI)	(4)Log(HHI)	(5)Log(HHI)	(6)Log(HHI)	(7)Log(HHI)
Year Trend	0.00420*** (0.000148)	0.00423*** (0.000168)	0.00268*** (0.000158)	0.00177*** (0.000140)	0.000760*** (0.000203)	0.000565** (0.000256)	0.00143*** (0.000504)
log(Emp)			-0.226*** (0.000824)	-0.213*** (0.000832)			
Constant	-9.458*** (0.298)	-9.520*** (0.338)	-5.810*** (0.318)	-4.022*** (0.281)	-2.551*** (0.407)	-2.176*** (0.513)	-4.214*** (1.013)
R^2	0.001	0.001	0.123	0.397	0.901	0.898	0.871
Mean Log(HHI)	-1.02	-1.02	-1.02	-1.02	-1.02	-1.04	-1.34
NAICS FE				Yes			
Estab FE					Yes	Yes	Yes
N	3,499,462	2,913,334	2,913,334	2,913,334	2,913,334	2,006,977	407,035

Observations weighted by survey weight. Columns (2)-(7) excludes the following NAICS sectors: 11, 21, 53, 61, 81, and 92.

Column (6) limits to multi-unit establishments. Column (7) limits to manufacturing sector.

Robust standard errors clustered at establishment level (UDBnum) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2 estimates the analogous regression specifications for another establishment-level outcome measure, the log number of occupations. The baseline specification in Column 1 shows that the number of 5-digit occupations per establishment fell by 14% from 1999 to 2016.

Table 2: Number of Occupations (1999 – 2016)

	(1)Log(#)	(2)Log(#)	(3)Log(#)	(4)Log(#)	(5)Log(#)	(6)Log(#)	(7)Log(#)
Year Trend	-0.00796*** (0.000174)	-0.00765*** (0.000198)	-0.00397*** (0.000153)	-0.00279*** (0.000135)	-0.00218*** (0.000213)	-0.00162*** (0.000269)	-0.00615*** (0.000481)
log(Emp)			0.536*** (0.000774)	0.522*** (0.000803)			
Constant	17.44*** (0.349)	16.83*** (0.398)	8.027*** (0.308)	5.701*** (0.272)	5.837*** (0.427)	4.701*** (0.541)	14.29*** (0.966)
R^2	0.002	0.002	0.475	0.634	0.928	0.925	0.934
Mean Log(#Occs)	1.45	1.46	1.46	1.46	1.46	1.44	1.93
NAICS FE				Yes			
Estab FE					Yes	Yes	Yes
N	3,499,462	2,913,334	2,913,334	2,913,334	2,913,334	2,006,977	407,035

Observations weighted by survey weight. Columns (2)-(7) excludes the following NAICS sectors: 11, 21, 53, 61, 81, and 92.

Column (6) limits to multi-unit establishments. Column (7) limits to manufacturing sector.

Robust standard errors clustered at establishment level (UDBnum) in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.2 Identifying Outsourceable Occupations

Next we attempt to identify occupations that are likely to have faced outsourcing throughout our study period, beyond the common occupations described by or focused on by previous studies. Out of over 800 detailed occupation codes, we limit our analysis to the top 50 ranking occupations in terms of employment in the 2000-2005 period. Next, we construct a measure that looks at overall employment in an occupation divided by the total number of establishments reporting at least one worker in that occupation. We call this the Employment to Establishment ratio, and calculate it for each occupation-by-survey wave. This measure captures the fact that when employers outsource a given occupation, they will no longer be considered an establishment reporting that occupation, and therefore the denominator of the E:E ratio shrinks. By occupation, we calculate the mean of this ratio for three separate time periods: 2000-2005, 2006-2010, and 2011-2016. Table 3 lists in order of employment the 20 occupations (out of the top 50) with the largest percentage change in Employment to Establishment ratio from the 2000-05 to 2011-2016 period.

However, increases in the E:E ratio may be partly driven by growth in labor demand for a given occupation, as evidenced by Registered Nurses and Home Health Aides. Therefore, we calculate an alternative measure of concentration named the Mean Proportional Deviation, defined below. The reasoning behind this measure is that it captures changes in the deviation of occupational employment shares across establishments. If an occupation is becoming increasingly concentrated in a few specialized employers, there will be a growing dispersion in employment shares, with firms that haven't outsourced employing low shares while the providers of outsourced labor employ large shares.

Table 3: Occupations with Largest Growth in Employment to Establishment Ratio

SOC Code	Description	Rank 00-05	Rank 11-16	E:E Ratio 00-05	E:E Ratio 06-10	E:E Ratio 11-16	% Change in E:E
29-1141	Registered Nurses	3	1	35.84	39.69	42.67	19%
53-7062	Laborers and Freight, Stock, and Material Movers, Hand	4	2	14.81	16.30	18.33	24%
51-2092	Team Assemblers	8	9	32.78	33.46	36.90	13%
37-2011	Janitors and Cleaners, Except Maids and Housekeeping Cleaners	9	11	6.96	8.19	8.49	22%
33-9032	Security Guards	12	13	20.46	21.68	23.07	13%
43-1011	First-Line Supervisors of Office and Administrative Support Workers	16	15	2.77	3.05	3.18	15%
53-7064	Packers and Packagers, Hand	19	23	15.84	17.38	19.94	26%
37-2012	Maids and Housekeeping Cleaners	24	24	12.68	14.49	15.15	19%
53-7051	Industrial Truck and Tractor Operators	25	34	10.34	11.02	12.16	18%
53-3033	Light Truck or Delivery Services Drivers	26	26	5.79	6.40	6.76	17%
31-1011	Home Health Aides	28	18	35.68	42.26	53.07	49%
13-2011	Accountants and Auditors	29	20	2.93	3.29	3.71	26%
43-4171	Receptionists and Information Clerks	30	35	2.63	2.71	2.97	13%
39-9021	Personal Care Aides	32	7	49.53	65.62	57.30	16%
35-2021	Food Preparation Workers	38	45	8.17	8.66	9.89	21%
41-9041	Telemarketers	44	74	23.96	24.84	31.36	31%
15-1132	Software Developers, Applications	46	25	14.40	14.68	16.01	11%
15-1121	Computer Systems Analysts	47	36	9.91	10.92	11.24	13%
51-4041	Machinists	49	48	9.22	10.26	10.53	14%
43-3021	Billing and Posting Clerks	50	49	3.21	3.67	4.16	30%

Notes: Listed are the 20 occupations within the Top 50 that have the highest growth in their E:E Ratios. Rank 00-05 denotes employment ranking of occupation over 2000-2005 period. Rank 11-16 denotes ranking of occupation over 2011-2016 period. E:E Ratio 00-05, 06-10, 11-16 denotes Employment to Establishment ratio over 2000-2005, 2006-2010, and 2011-2016 periods respectively. % Change in E:E denotes the percentage change of the E:E Ratio from 2000-2005 to 2011-2016.

Table 4: Occupations with Largest Growth in Mean Proportional Deviation

SOC Code	Description	Rank 00-05	Rank 11-16	MPDev 00-05	MPDev 06-10	MPDev 11-16	% Change in MPDev	% Change in E:E
37-2011	Janitors and Cleaners, Except Maids and Housekeeping Cleaners	9	11	1.49	1.63	1.73	16%	22%
43-1011	First-Line Supervisors of Office and Administrative Support Workers	16	15	1.73	1.83	1.98	15%	15%
37-2012	Maids and Housekeeping Cleaners	24	24	1.73	1.97	2.17	25%	19%
53-3033	Light Truck or Delivery Services Drivers	26	26	1.79	1.86	2.05	15%	17%
43-4171	Receptionists and Information Clerks	30	35	2.62	3.16	3.58	37%	13%
35-2021	Food Preparation Workers	38	45	1.69	2.13	2.39	42%	21%
43-3021	Billing and Posting Clerks	50	49	2.09	2.49	2.75	32%	30%

Notes: Listed are the occupations from the previous table that have the highest growth in their MPDev. Rank 00-05 denotes employment ranking of occupation over 2000-2005 period. Rank 11-16 denotes ranking of occupation over 2011-2016 period. MPDev 00-05, 06-10, 11-16 denotes Mean Proportional Deviation over 2000-2005, 2006-2010, and 2011-2016 periods respectively. %Change in MPDev denotes the percentage change of the MPDev from 2000-2005 to 2011-2016. % Change in E:E denotes the percentage change of the E:E Ratio from 2000-2005 to 2011-2016.

$$MPDev_j = \frac{1}{n} \sum_{i=1}^n \left| \frac{share_{ij} - avgshare_j}{avgshare_j} \right|$$

$$where \quad share_{ij} = \frac{Emp_{ij}}{TotEmp_i} \quad and \quad avgshare_j = \frac{\sum_{i=1}^n Emp_{ij}}{\sum_{i=1}^n TotEmp_i}$$

We can then limit the above list of occupations to a set of occupations that show large changes (over 15%) in their Mean Proportional Deviation from 2000-05 to 2011-16. We claim that these occupations are likely to have been subjected to greater outsourcing from 2000 to 2016, and warrant further study.

4 Industry Concentration

Our results on changes in the industrial concentration of firm activity over time are pending disclosure avoidance review by the U.S. Census Bureau, and will be available shortly in subsequent drafts.

5 Conclusions

We use the Bureau of Labor Statistics and Census data to investigate the concentration of U.S. firms' occupations and industries over time. We find strong evidence for increased concentration of firms in terms of (a) decreases in the number of occupations and industries in which firms are active (b) increases in the share of employment accounted for by firms' most common occupations and most important industries (c) rising occupational HHI and industrial HHI indices. The mean number of 5-digit occupations per establishment has fallen from 6.5 to 5.5 since 2000, with the top-3 occupations now accounting for over 85% of total establishment employment. The concentration of activities within firms which we have documented is likely linked to the rise of outsourcing and pressure on firms to focus on their core competencies. In future work we hope to explore this further, as well as the potential role of increased firm concentration on the rising segregation of employees by income and education across firms.

References

Fort, Teresa C. and Shawn D. Klimek (2018). “The Effects of Industry Classification Changes on U.S. Employment Composition,” *Center for Economic Studies Working Paper 2018-28*, U.S. Census Bureau.

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

Prahalad, C.K. and Gary Hamel (1990). “The Core Competence of the Corporation,” *Harvard Business Review*.

Song, Jae, David Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter (2019). “Firming Up Inequality,” *Quarterly Journal of Economics*.

Weil, David. “The Fissured Workplace,” Harvard University Press, Cambridge, MA, 2014.