6-2019

Industry Concentration and Information Technology

James Bessen

BU School of Law

Industry Concentration and Information Technology

Boston University School of Law
Law & Economics Paper No. 17-41

James Bessen
Boston University School of Law

Industry Concentration and Information Technology

Author: James Bessen
jbessen@gmail.com
207-607-1334

Technology & Policy Research Initiative
Boston University School of Law
765 Commonwealth Ave.
Boston, MA 02215

Industry Concentration and Information Technology

By James Bessen

Technology & Policy Research Initiative
Boston University School of Law

June 2019

Abstract: Industry concentration has been rising in the US since 1980. Does this signal declining competition and need for a new antitrust policy? Or are other factors causing concentration to rise? This paper explores the role of proprietary information technology (IT), which could increase the productivity of top firms relative to others and raise their market share. Instrumental variable estimates find a strong link between proprietary IT and rising industry concentration, accounting for much of its growth. Moreover, the top four firms in each industry benefit disproportionately. Large investments in proprietary software—$250 billion per year—appear to significantly impact industry structure.

Keywords: industry concentration, antitrust, information technology, computers, productivity dispersion

JEL codes: D4, O33, L10, L4

Thanks for helpful comments from David Autor, Maarten Goos, Joe Mazur, Mike Meurer, Nancy Rose, Ronja Rottger, Anna Salomons, Mike Scherer, Dick Schmalensee, Rob Seamans, Carl Shapiro, Tim Simcoe, John Turner, Gabriel Unger, Jeroen van den Bosch, Hal Varian and participants at the IIOC, the TPRI seminar Competition Conference, and the NBER Productivity seminar. Thanks to IBM for financial support during this research.

Industry concentration at the national level has been rising across sectors in the US since the 1980s. Autor et al. (2017) find that from 1982 to 2012 the share of shipments made by the top four firms in four-digit manufacturing industries grew 4.5% in the US with similar increases in most other major sectors.[1] At the same time, evidence shows a concomitant rise in profit margins and markups in the US (Rognlie 2015, Barkai 2016, de Loecker and Eeckhout 2017). Figure 1 shows the recent rise in operating margins. What is driving this change and what is its significance?

Some see rising concentration as a sign of decreasing competition that might lead to higher prices, less innovation, and greater wage inequality.[2] Grullon et al. (2016) attribute the rise in industry concentration partly to lax antitrust enforcement of mergers and acquisitions. Gutierrez and Philippon (2017, 2018) suggest that growing federal regulation, weakened antitrust, and corporate lobbying might be reducing competition specifically in the US. If these views are right, then perhaps antitrust enforcement needs to be strengthened or other policy changes made to increase competition.

However, rising industry concentration does not necessarily imply declining competition. As Demsetz (1973) argued, concentration can also rise when some firms grow faster because they are more efficient. In this case, rising concentration would reflect greater innovation and social benefit. The policy implications from rising industry concentration depend very much on what is *causing* the increase.

---

[1] See also White and Yang (2017) on trends in aggregate concentration. Rinz (2018) and Rossi-Hansberg, Sarte, and Trachter (2018) find that *local* concentration ratios have been falling.

[2] The Economist, "Too much of a good thing," March 26, 2016. National markets identified in the Economic Census do not correspond to the "relevant markets" used in antitrust analysis, however, the general rise in national concentration ratios might reflect important changes nevertheless.

Declining competition is one possible cause, but there are others. Increased exposure to global competition could increase the market share of the most productive firms and force less efficient producers to drop out (Mellitz 2003). Another possibility is that some firms—but not all—benefit significantly from new technologies. Thanks to new technology, these firms earn higher profits and realize larger market share, hence higher concentration. In some markets, notably in some high-tech industries, network effects may provide substantial benefits to the largest players, creating "winner-take-all-markets" (Autor et al. 2017). In other industries, technology might boost the market shares of some firms if there are economies of scale or if the technology is not accessible to all firms. By these mechanisms, rising concentration could signal growing productivity dispersion rather than a decline in competition.

This alternative view is bolstered by several studies that point to a growing divergence in firm productivity within industries—the gap between the top performing firms and the rest is growing (Andrews et al. 2016; Berlingieri et al. 2017). Peltzman (2018) finds that within manufacturing, growing concentration is associated with higher productivity. In addition, rising productivity gaps and rising markups are observed across developed economies (de Loecker and Eeckhout 2018, Diez, Leigh, and Tambunlertchai 2018, Calligaris, Criscuolo, and Margolin 2018), undercutting the notion that specific US domestic policies are the main causal culprit.

To understand the significance of rising concentration, it is necessary to disentangle the factors that are causing concentration to rise across industries. This paper explores the role of one major factor: the large investments that firms are making in proprietary information technology (IT). According to BEA estimates, in 2016, firms invested $250 billion in proprietary software development (self-developed and contracted). That is nearly as much as all private nonresidential investment in equipment and structures, net of depreciation. Furthermore, large firms make disproportionately larger investments in developing software, perhaps because of scale economies (Unger 2019). Hence, large investments in proprietary IT might allow big firms to grow faster, providing a possible explanation for rising industry concentration.

This paper explores the link between national industrial concentration ratios with proprietary IT measured as the share of software developers in the industry workforce. A simple binned scatterplot in Figure 2 shows a correlation between these two variables. Of course, proprietary software might be endogenous. For example, large firms in concentrated industries might need greater IT resources to manage their enterprises. To achieve identification, I instrument industry IT intensity with a measure of the share of jobs in each industry that are sedentary. The motivation for this instrument is that computers are more readily adopted in sedentary occupations yet industry concentration is not likely to influence the sedentariness of occupations as measured in 1977, the source year of the data. Placebo tests provide some support for this assumption. In addition, I use two other instruments that should be independent of changes in US competition policy, especially since 1980: the IT share of the workforce in 1980 and the IT share of investment in 18 European countries.

The main contribution of this paper is that industry use of proprietary IT is associated with higher industry concentration ratios (shares of sales to the top firms) and with more rapid growth in concentration ratios. The effect is large—it accounts for most of the observed rise in concentration ratios—and the instrumental variable analysis provides some evidence that the relationship is causal. In contrast, measures of merger and acquisition activity are not positively associated with changes in concentration. In addition, industry use of proprietary IT is associated with larger revenues per establishment and higher labor productivity among the top four firms within each industry, both in absolute terms and relative to other firms in the industry. While these findings by themselves provide little support for a change in antitrust policy, they do indicate that large and rising investment in proprietary IT systems is affecting industry structure and is an important phenomenon to study. Moreover, these changes are occurring across all sectors; they are not just about Big Tech.

Background

Rapidly falling prices for computer hardware and strong growth of pre-packaged software have suggested to some that IT may be "levelling the playing field," allowing small firms technology to compete with larger rivals.[3] However, IT investment, especially at large firms, has become dominated by proprietary technology. The majority of firm IT investment today goes to custom systems for the firm's own use, either developed in-house or by contractors. According to BEA statistics from 2016, custom plus own-account software account for 55% ($250 billion, up from 33% in 1985) of the total private investment in software, computers, and peripherals ($452 billion). A key difference is that these systems, as opposed to off-the-shelf products, can deliver competitive advantage. For example, since the 1970s, off-the-shelf barcode scanners and associated computer programs have been available to retail stores both large and small.

[3] See, for instance, Sarah Schafer, "How Information Technology Is Leveling the Playing Field," *Inc.* (1995).

These systems provide proven productivity advantages. However, these advantages are not large and, by itself, it is unlikely that the barcode scanner increased concentration in retail industries.

But Walmart integrated these scanners into a complex proprietary system. In 1990, Walmart introduced a system that linked suppliers to stores to headquarters, providing suppliers detailed inventory data for each store. The technology, combined with complementary changes in the organization of distribution centers and stores, allowed Walmart to adjust rapidly to changes in demand, for instance, to identify hot-selling items and to get them on store shelves quickly. The system speeded the delivery of goods, reduced inventory requirements, increased the number and variety of items sold in each store, reduced prices, and delivered dramatically faster productivity growth. Few rivals could match Walmart's technology. Basker (2007) suggests that Walmart alone accounts for most of the growth in productivity in general merchandise retailing from 1982 to 2002 and this explains its growing market share. In 1982, Walmart accounted for 3% of the sales of US general merchandise retailers; thirty years later, Walmart's US sales comprised 52% of industry sales.

Moreover, investments in proprietary IT are being made across all major sectors, not just by big tech companies nor by just a few companies like Walmart. Big banks developed IT systems to handle credit card operations; Boeing developed systems to design large aircraft.

Proprietary IT can contribute to rising industry concentration in multiple ways (not mutually exclusive). These large systems may create substantial economies of scale, allowing large firms to grow faster. Below I show evidence that IT is associated with larger establishment size, especially at leading firms. While general purpose IT might also exhibit scale economies, firm behavior suggests scale effects are tightly linked to proprietary software as opposed to off-the-shelf software. Pre-packaged software as a share of investment declines sharply with firm size while own-developed software increases, dramatically so for the largest firms (Unger 2019). Software developers comprise 4.1% of the workforce at firms over 1,000 employees but only 1.3% of the workforce in firms with 50 or fewer employees.[4]

Proprietary IT systems may also create persistent productivity advantages; they may include innovations that increase productivity by decreasing costs, improving service quality, or allowing targeting or price discrimination.[5] These advantages will not be available to rival firms if they are protected by trade secrecy or patents or if they depend on complementary knowledge and skills of managers or developers. In this case, the proprietary IT generates innovation quasi-rents. Below I show evidence that IT is associated with greater labor productivity, especially at leading firms (see also Akcigit and Ates 2019; Calligaris, Criscuolo, and Margolin 2018; and Dunne et al. 2004). In many IO models, firms that increase their productivity will also increase their market share and empirical studies find strong support for this association (Decker et al. 2018). If larger firms tend to invest relatively more in proprietary software, they will tend to become more productive, to grow faster, and industry concentration will rise.

Large IT investments may constitute a fixed cost that serves as a barrier to entry. In particular, large IT investments may constitute an "endogenous fixed cost" in the sense

[4] Using data from the Current Population Survey, 2010 to 2017.

[5] Thanks to an anonymous referee for pointing out the potential role of price discrimination.

of Shaked and Sutton (1983, 1987), leading to a “natural” oligopolistic industry structure. For instance, Ellickson (2006) finds evidence that retail industries are such natural oligopolies and Crouzet and Eberly (2018) attribute the growth in retail industry concentration to rising investment in intangibles, including IT.

Several papers are related to this one. Brynjolfsson et al. (2008) find that all industries exhibit growth in concentration from 1996-2006 but that IT intensive industries show somewhat faster growth on average during this period.[6] The present paper goes beyond this by using a more detailed set of industries, using instrumental variables, and performing a supplementary analysis on differences between the top firms and the rest within each industry. Finally, Tambe and Hitt (2012) and Harrigan et al. (2016) also use the employment share of IT workers as an independent variable to explore firm productivity and job polarization respectively.

Data

Industrial concentration

The concentration data come from the quinquennial Economic Census reports that use the NAICS industry classification, beginning in 1997 through 2012. The Census reports the share of industry revenues (or shipments) going to the top 4, 8, 20, and 50 firms in each NAICS industry at the 2, 3, 4, 5, and 6 digit levels. In addition, it reports the number of establishments, annual payroll, and number of employees for the industry as a whole and for the top firms within the industry (the latter data are missing for

[6] Their measure of concentration is a Herfindahl index based on Compustat data.

manufacturing industries). I also use data from the 1977 Economic Census for the manufacturing sector, using a walkway to convert SIC industries to NAICS (see below).

Census industry definitions, even at the 6 digit level, do not necessarily correspond to the market definitions needed for competition analysis (Shapiro 2017). For example, the airline industry shows increased concentration by these measures, but detailed analysis of the number of competitors for different routes shows that competition at the route-level has not declined. Moreover, rising concentration at the national level appears to be accompanied by increased competition at the local level (Rinz 2018; Rossi-Hansberg, Sarte, and Trachter 2018). Nevertheless, rising concentration ratios from the Economic Census have been used to argue that competition is decreasing and, in any case, they do signal an important trend that *something* affecting industry structure is changing, even if it is not the level competition.

Note that I exclude some industries where software is a major part of their products, for reasons related to the IT variable discussed below. While some of the public concern about competition has focused on large tech firms, the focus here is on the many industries in diverse sectors experiencing rising concentration. Large tech firms might have special characteristics, such as network effects, that raise distinct concerns not shared in other sectors.

The Economic Census data have the advantage that they count all firms and establishments in each industry. Some studies have used concentration ratios computed for publicly firms listed in Compustat (Grullon et al. 2016; Guttierez and Philippon 2017). Those data have the advantage of being available annually and for a longer period of time. But they also have some disadvantages: Compustat typically reports worldwide sales, not domestic sales, and the sample excludes private firms. If we want to analyze concentration in domestic markets, it can be misleading to use measures based on international sales. And it appears that private firms make a large difference. The Compustat concentration ratios are only weakly correlated with the ratios provided by the Economic Census.[7] To avoid conflating issues about concentration with issues about firms' changing preferences about being publicly listed and firms' changing international exposure, I employ the Economic Census data.

Proprietary IT

The paper seeks to capture the extent to which firms use proprietary IT systems. Firms building proprietary systems will typically hire software developers and systems analysts to design, build, and maintain these systems. General computer use for common office applications does not require such personnel. Proprietary systems might incorporate off-the-shelf components including software (e.g., SAP software), but these components are bundled with firm-specific software.

For each industry, I measure proprietary IT as the software share of the workforce, specifically, the share of hours worked by IT personnel, identified as people in the following occupations: computer systems analysts and computer scientists, operations and systems researchers and analysts, and computer software developers.[8]

---

[7] I ran several tests. For example, I calculated the Compustat four-firm concentration ratios for 2012 for three-digit NAICS industries. The correlation coefficient between these data and the corresponding four-firm ratios from the Economic Census was 0.196.

[8] Hours worked is calculated as weeks worked last year time usual hours worked per week times the person weight. For 2012, weeks worked is intervalled; I assign a numeric value based on the means for 2007. Note that these occupations comprise about 83% of all IT employees excluding managers.

Since the aim is to measure the use of custom proprietary IT, I exclude industries that are involved in creating information technology products.[9] These industries employ IT personnel in designing and producing products, not just in building systems for their internal use. Also, to reduce measurement error in small industries, the sample excludes the smallest 5% of industries by employment.[10]

Some proprietary IT is contracted rather than developed inhouse. I assume that firms building proprietary systems will typically hire software developers and systems analysts to design, build, and maintain these systems even if much of the work is done by outside contractors. In fact, the software share of the workforce is correlated with BEA software investment measures that do include contracted software.[11] Tambe and Hitt (2012) find that a similar labor-based measure corresponds with a variety of other measures of IT.

Data on the workforce come from the public use samples of the American Community Surveys (Ruggles et al. 2015). These data are not available for 1997, so the some of the analysis is restricted to 2002, 2007, and 2012.[12] The American Community

---

[9] These include NAICS 5112, software publishers, 5181, Internet service providers and web search portals, 5182, Data Processing, Hosting, and Related Services, 5191 Other information services, 5415 Computer Systems Design and Related Services, 3341 Computer and peripheral equipment manufacturing, 3342 Communications Equipment Manufacturing, 3344 Semiconductor and Other Electronic Component Manufacturing, and 3345 Navigational, measuring, electromedical, and control instruments manufacturing.

[10] That is, it excludes industries with fewer than 28,748 employees.

[11] The BEA/BLS Integrated GDP-Productivity accounts report the capital income of software investment by year for 61 private industries (see https://www.bea.gov/industry/an2.htm#integrated). I aggregated my data up to the BEA/BLS industries (my data have nearly four times as many industries) and compared the share of IT workers in the industry workforce to the share of software compensation in total gross output. The association was highly significant with a correlation coefficient of .42.

[12] While workforce data is available for other sources for 1997, such as the Current Population Survey, the sample sizes of these sources are far smaller than those of the ACS, making detailed industry analysis infeasible.

Surveys use modified NAICS industry codes which are aggregated to different levels. Some industries are identified at the 6-digit level while others are only identified at the 3-digit level. I match these industries to the corresponding industries in the Economic Census to obtain a sample of 730 industry-year observations over three years at different (non-overlapping) levels of industry classification.[13]

I also use data for the manufacturing sector for 1977, using the 1980 Census of Population to obtain measures of the software share per industry. To make the 1977 Economic Census data comparable both to the Census of Population and to the later Economic Censuses, I match the 1977 industries. Where the target data use a higher level of industry aggregation, I averaged the 1977 industry data on concentration, weighting by shipments per detailed industry.

Operating Margins

As a robustness check, I also look at the relationship between proprietary IT and the growth of firm operating margins. For this analysis, the main sample consists of Compustat firms traded on US exchanges in 2000 and 2014, excluding financial firms, matched to industry IT systems data, totaling 1,532 firms. I exclude firms that are missing data on market value, sales, and assets, firms where R&D exceeds half of revenues (startup mode), and I exclude the 5 percent tails of the dependent variable (operating margin, that is, operating income after depreciation before taxes, R&D, and

[13] There are 75 3-digit industries, 459 4-digit, 151 5-digit, and 45 6-digit industries. Note that there are some minor changes in the NAICS classification between 2002 and 2012, so that some industries are not reported for all three years.

advertising expense all divided by revenues) to counter measurement error at the extremes. I use the method of Lewellen and Badrinath (1997) with the NIPA investment deflator to calculate the net capital stocks. Stocks of R&D and advertising and marketing expenditures are computed using the perpetual inventory method.[14] Industry level IT capital is also calculated using the perpetual inventory method where annual investment consists of the deflated wages paid to IT personnel in the industry.[15] As a control in the operating margin regressions, I use a measure of industry regulation developed by Al-Ubaydli and McLaughlin (2015) that is based on an industry-relevance weighted count of words in the Code of Federal Regulations.[16]

Summary statistics

Table 1 provides some summary statistics on the sample of industries. On average, IT workers account for 2.2% of hours worked. The table shows the four different concentration ratios. Relatively few industries could be described as monopolies or oligopolies; the top four firms account for the majority of revenues in only 15% of the industries. But industries have been growing more concentrated. The table shows the mean five-year change in concentration ratios from 1997 to 2007, before the recession; the mean changes from 2007 to 2012 were smaller. Note that most of the increase in

---

[14] The R&D stock is calculated assuming a 15% annual depreciation rate and an 8% pre-sample growth rate (Hall 1990); R&D expenditures are deflated using an R&D deflator. The advertising stock is based on advertising and marketing expenditures and assumes a 45% annual depreciation rate and 5% pre-sample growth rate (Villalonga 2004, p. 217).

[15] I assume a 15% depreciation rate and a 2% pre-sample growth rate based on the average growth rate from 2000-2014. I divide the IT capital by the number of workers in each industry each year to obtain a scaled measure of IT capital per worker.

[16] Al-Ubaydli and McLaughlin use an algorithm to probabilistically assign each section of the Code to a specific NAICS industry. They do this assignment for sets of 2-digit, 3-digit, and 4-digit NAICS industries. The result is a time series of the extent of regulation for specific industries since 1970.

concentration can be attributed to the growing share of the top four firms; the increase in the share of the top 50 firms is not much larger than the increase for the top four. Also, consistent with prior literature (Schmalensee 1989), the top firms in each industry tend to have larger plants (revenues / establishment), higher labor productivity (revenues / employee), higher pay, but lower labor share of output.

Table A1 in the Appendix displays the distribution of observations across industry sectors, defined as the first digit of the industry NAICS code. It also displays the average change in the four-firm concentration ratio for each sector from 2002 to 2007. Most sectors show rising concentration, except for education and health, which have a high nonprofit component.

Instrumental variables

Firm investments in information technology might be endogenous, reflecting other factors that could also be related to industry concentration. This might confound the analysis of the impact of IT on concentration or the analysis of operating margins. For example, faster growing firms might invest more in IT in order to manage their more rapid growth; they would become larger, possibly increasing industry concentration, and their growth would be correlated with IT. But in this case, growth in market share would cause IT spending rather than the reverse.

In order to correct for reverse causality and other confounding influences in the analysis of industry concentration and operating margins, I estimate the relationships using three different instrumental variables. The ideal instrument should be correlated with (but independent of) IT and it would also plausibly satisfy the exclusion restriction; that is, the ideal instrument would not influence industry concentration except through IT.

To instrument the software share of hours, my main instrumental variable is a measure of industry sedentariness derived from the Dictionary of Occupational Titles (1977). The notion here is that it is easier to implement computer technology in industries with more sedentary employees because seated employees can more advantageously use desktop computers or terminals. These industries should therefore tend to adopt IT somewhat earlier and somewhat more intensively, all else equal.

I use a measure of sedentariness derived from the 1977 edition of the Dictionary of Occupational Titles; this was before most occupations used computers so that computers likely had little effect on the sedentariness of occupations. The US Department of Labor has sought to define aspects of some 14,000 distinct jobs, including a measure of how sedentary the job is, publishing the fourth edition of this work in 1977.[17] England and Kilbourne (2013) have mapped the DOT occupations to Census detailed occupation codes, averaging them to this higher level of aggregation. Using these occupations, I calculate the distribution of sedentary occupations across industries using the 1980 Census public use sample.

In order to use sedentariness as an instrument, I need to map it to the same industry categories used for the dependent variable, industrial concentration. For the analysis of concentration from 1997 through 2012, I develop a walkway to map the 1980 Census industries to the NAICS categories used in the Economic Censuses, using the most disaggregated classifications possible. These industry categories, however, differ

---

[17] The DOT reports a job characteristic called STRENGTH, which rates the physical demands of the job on a scale of 1, for sedentary occupations, to 5, for very heavy work. Only the first category relates to sedentariness; the other categories relate to level of exertion required. Since the England and Kilbourne data report averages for an occupation, I flagged an occupation as being sedentary if its STRENGTH rating is less than 2.

from those used in the American Community Survey (ACS) used to derived the measures the software share of the workforce. Moreover, the ACS was not conducted in 1997. For this reason, rather than do a two-stage least squares for the years 1997-2012, I do a reduced form IV estimation, directly regressing industry concentration on the instrumental variable, sedentariness.[18]

To use a reduced form IV for the disaggregated data from 1997 through 2012, the instrument should be correlated with the endogenous variable. Table 2 shows correlation coefficients and first stage regressions for those industries where both the sedentariness instrument (from the DOT and 1980 Census) and the measure of the software share of the workforce (from the ACS) are available. The correlation coefficients for the years 2002, 2007, and 2012 range from .307 to .328 and the regression coefficients are highly significant. One concern is that the rise of mobile computing might correspond to a weakening of the instrument, which is based conceptually on desktop computing. While the regression coefficient on the sedentariness variable did decline somewhat from 2002, this difference is not statistically significant and the correlations and regression R-squared statistics did not weaken.

Sedentariness and computer use vary substantially across sectors. Table A2 in the Appendix shows the mean sedentariness of each 1-digit NAICS sector as well as the index for the lowest and highest industry within each sector. Finance, real estate, and business services is the most sedentary sector (mean .70) while agriculture is the least sedentary (mean .14). However, the differences in the sedentariness index between the

---

[18] For the analysis from 1977 to 2002, I aggregated the data to industry categories that correspond to the ACS, so a full two-stage least squares is possible. Aggregation dilutes the concentration measures, so a disaggregated approach is preferred for the main analysis.

low and high industries within each sector show that there is significant variation in the index within sectors. For example, within manufacturing, Animal Slaughtering and Processing has a sedentariness index of .12, but Aerospace Products and Parts has a sedentariness index of .73. Moreover, the correlation between sedentariness and IT share of the workforce, estimated for 2002, 2007, and 2012, is substantial for all sectors except for Other Services. Thus, the link between this instrumental variable and the software share is not mainly driven by a few industries or sectors.

One concern is that sedentariness might be linked to other occupational characteristics that somehow affect industry concentration. Specifically, while sedentary occupations are more likely to use computers, they are also more likely to handle paper documents. Sedentariness is likely correlated with the use of desks, paper, and pencils. Dinardo and Pischke (1997) famously found that pencil use is correlated with higher wages, likely reflecting unobserved worker characteristics of those workers who select into pencil-using occupations. Sedentariness might well be correlated with such characteristics and also with higher wages.

These correlated variables might cause a problem for the instrument if they were *also* correlated with the outcome variable, industry concentration. Evidence in Table 3 suggests that this second correlation is not a significant problem. This table regresses several measures of industry concentration and the growth in industry concentration against three industry characteristics: the share of workers in professional and managerial occupations, the mean years of schooling of workers in the industry, and the mean log industry wage. The regressions also include dummy variables for year, industry sector, and the number of digits in the industry classification, as are used in the regressions on industry concentration below. Joint tests of the significance of these variables cannot reject the null hypothesis that they are all zero. Individually, the coefficients are not statistically significant except for weak significance (10% level) of the wage variable in the two broadest measures of industry concentration. These estimates appear to rule out the possibility that the correlation between sedentariness and industry concentration spuriously reflects the effect of professional/managerial work, education, or wages.

Further evidence in support of the validity of the exclusion restriction comes from placebo tests. The left side of Table 4 reports regressions on industry concentration using data from the 1977 Economic Census for the manufacturing sector and also from the Economic Censuses of 1997, 2002, 2007, and 2012. The regressions show that the instrumental variable is not significantly correlated with the four-firm concentration ratio in 1977, but the association is highly significant for the more recent sample of manufacturing industries. The assumption in this paper is that the correlation during the recent period reflects the greater use of information technology since 1977. A similar pattern is seen in the right panel of the table which regresses firm operating margins on the instrumental variable with various controls corresponding to the analysis below. Again, the coefficient for 1977 is not significant while the coefficient for the recent period is highly significant.

This finding does not definitively eliminate the possibility that some third factor could be responsible for a spurious link between proprietary IT use and industry concentration or operating margins. However, it does mean that such a third factor could not have had significant influence prior to 1980 and its influence must have grown more or less concurrently with the rapid growth in IT systems use after 1980.

To bolster the validity of this IV analysis, I also use two other instruments directed to the particular concern that rising concentration might reflect changes in US competition policy, especially after the 1980s. The first supplementary instrument uses the share of software developers in each industry's workforce using data from the 1980 Census Public Use sample. This should be independent of subsequent policy changes.

The second additional instrument measures the share of software investment in total investment of the industries of 18 European countries obtained from the EU-KLEMS database (Jäger 2017).[19] These data are grouped into far fewer industrial categories (24 that match the Economic Census), so for each European industry I calculate a weighted average (by shipments/revenues) of the US industry concentration and software share variables. To the extent that competition policy differs between the US and Europe, this instrument should be independent of US policy yet still be correlated with US IT use. Industry concentration in Europe reflects distinct factors such as the formation of EU common markets. Empirical studies differ as to whether industry concentration is rising or falling in Europe since 2000, but competition policy is seen to differ significantly (Gutierrez and Philippon 2017; Bajgar et al. 2019). Both of these supplementary instruments should be independent of US policy since the 1980s although they might be correlated with some third factor associated with industry concentration other than IT.

[19] These countries, determined by data availability, are Austria, Cyprus, Czechoslovakia, Germany, Denmark, Estonia, Greece, Spain, Finland, France, Hungary, Italy, Luxembourg, Latvia, Netherlands, Portugal, Sweden, and Slovenia.

Empirical Findings

Basic regressions on concentration ratios

Table 5 shows basic regressions on the different concentration ratios. The regression estimates concentration ratio j for industry i during year t:

$$C_{ijt} = \beta \cdot IT_{it} + \alpha_t + \delta_I + \gamma_n + \epsilon_{it}$$

where IT_{it} is the measure if proprietary IT use, δ_I is a dummy variable for industry sector (1-digit NAICS code), and γ_n is a dummy variable for the number of digits in the industry definition. The latter dummy variable is included because more narrowly defined industries are likely to have higher concentration ratios, all else equal. Table A3 in the Appendix breaks out the regression for the 4-firm concentration ratio by different industry digit levels. All show an association between IT share and industry concentration, but the estimates for more narrowly defined industries are larger and have greater statistical significance.

The top panel of Table 5 shows OLS regressions on the pooled (2002-2012) level of each concentration ratio with heteroskedasticity-robust standard errors without the industry sector dummy, δ_I. The coefficient of the share of IT workers in the workforce is significant for all concentration ratios. It is also economically significant. The sample mean of the software share of hours worked is 2.2%. At this mean, the software share is associated with an increase in the revenue share of the top four firm of 2.2% x 2.14 = 4.7%. This is comparable to the increase in four-firm concentration ratios reported by Autor et al. (2017) for most sectors since 1982. Since the share of IT workers was much smaller in 1982, proprietary IT use appears to "explain" most of the increase in industry concentration since then, loosely speaking.[20]

Since the panel is largely cross-sectional—the time dimension is at most 3 observations—estimates with full industry fixed effects may not be consistent. Adding industry sector fixed effects, δ_I, provides a degree of control for omitted variables associated with industry characteristics. Panel B shows these estimates. Generally, coefficient estimates and standard errors are slightly smaller. The within R-squareds are substantial, suggesting that even in this short panel, time variation provides significant identification.

One concern is that these estimates are unrepresentative because the sample does not accurately reflect business activity. The industries defined by the Census vary substantially by size. Panel C repeats the analysis of Panel B, but weighting observations by industry shipments/revenues. The coefficients are still all highly significant, but some decline and some increase. The weighted regression should also reduce measurement error in the software share--some small industries likely suffer from sampling variance because of limited data in the ACS.

Another concern with these estimates is the possibility that proprietary IT use might be endogenously related to the error term. Panel D reports the same regressions using the instrumental variable in place of the measure of software share of the workforce in a reduced form IV. The coefficients on sedentariness are all highly significant. To compare these estimates to the OLS estimates, it is necessary to scale them. I estimate a

[20] One concern is that many firms in education and health care are nonprofit, perhaps biasing the results. Repeating these regressions but excluding those industries (results not shown) makes little difference in the coefficients.

scaling factor by regressing the software share of the workforce on sedentariness with controls for year and sector for those industries where both data items are available. The scaling coefficient is 6.24. The bottom row of the panel displays the scaled regression coefficients. The estimates are somewhat higher than the OLS estimates.

The levels of industry concentration observed in the pooled sample roughly capture the increase in concentration brought about by the adoption of proprietary IT, occurring mainly since 1980 or so. A further test is to see whether IT is also related to the *growth* in concentration occurring during the sample period. Panel E makes reduced form IV estimates of the change in concentration ratios between 1997 and 2007. I exclude changes after 2007 because of possible confounding effects of the recession. The coefficients on sedentariness are significant, at least marginally, for three of concentration measures. The panel also shows scaled coefficient estimates. At the sample mean, the software share is associated with an increase in the four-firm concentration ratio of 0.56 x 2.2% = 1.2%. This is slightly smaller than the actual change in the mean four firm concentration ratio shown in Table 1, 1.43%.

To further bolster the analysis, Table 6 shows results using alternative instruments. The top panel uses the 1980 software share of the workforce in a reduced form IV estimation. The coefficients are all highly significant and the scaled coefficients are similar to those in Panel D of Table 5. The second panel shows a full two-stage least squares estimation using the aggregated industry categories of the EUKLEMS dataset. Here the coefficients are all significant at the 5% level; they are smaller, but that is not surprising given that the industries are more highly aggregated.

Finally, note that most of the increase in concentration ratios associated with IT is driven by the top four firms. That is, the coefficient for the eight-firm ratio is only slightly larger than the one for the four-firm ratio, implying that the market shares of firms five through eight grew relatively little. Similarly, for the other concentration ratios. For this reason, the remainder of the paper focuses on just the role of the top four firms.

Long differences

Table 7 extends this analysis by looking at the change in the four-firm concentration ratio from 1977 to 2002. This sample is for the manufacturing sector only, due to limitations in the available public data.[21] The first column uses the 1980 estimate of the software share and the second column measures the difference between the software shares in 1980 and 2002. The third column repeats the regression of column 1 using IV estimations. In all of these regressions, the coefficient on the software share is significant. The bottom of the table shows the sample means of the IT measures and product of these means and the software share coefficient. In each estimation, the software share accounts for a 3 – 5% rise in industry concentration, roughly corresponding to the actual increase found by Autor et al. (2017). In other words, IT use appears to account for much of the rise in industry concentration.

Other variables

Of course, other factors likely affect industry concentration in some sectors as well. These factors might confound the analysis if they are correlated with proprietary IT

[21] The sample also excludes industries where software development is part of the product and it excludes the 1% tails in the dependent variable (1 observation each) in order to limit measurement error.

use and also with industry concentration. Table 8 considers some possibly confounding variables: the number of establishments, merger and acquisition activity, exposure to imports, and industry growth. Including these variables in regressions along with the measure of proprietary IT use provides a robustness check on the IT coefficient.

Column 1 includes the number of industry establishments. The more establishments in an industry, the harder it might be for a few firms to capture a large market share. Also, rising entry barriers would tend to reduce the number of establishments, driving concentration up. Including this variable does not significantly change the coefficient on proprietary IT use and the coefficient on the number of establishments is weakly significant (P = .092), negative, and small. A supplementary regression (not shown) on the change in industry concentration from 2002 to 2007 against the change in industry establishments shows no significant relationship. Thus, the number of establishments does not confound the IT relationship.

Column 2 includes a measure of merger and acquisition activity. Grullon et al. (2017) argue that mergers and acquisitions are a major reason industry concentration is rising, which they attribute to lax antitrust enforcement. To measure industry M&A activity, I use data from Thomson Reuters SDC database of M&A transactions. Since acquisitions by large firms are those most likely to affect industry concentration and since large firms are more likely to be publicly listed, I extracted those acquisitions made by publicly listed firms. Excluding transactions where the acquirer did not obtain majority ownership or where ownership percentage was not reported, I matched these data with Compustat data for publicly listed firms, resulting in a list of 33,942 acquisitions by publicly listed firms from 1985 through 2001.[22] I use these data to construct an index of industry M&A activity prior to 2002. Using the Compustat historical NAICS assignments for each firm, I tabulated the number of acquisitions and the number of active publicly listed firms for each industry. I then calculated the index of M&A activity as the aggregate number of acquisitions per public firm for each industry over the entire period. The regression finds a negative coefficient on M&A activity that is not statistically different from zero. The coefficient on proprietary IT use changes only slightly. Using this measure, mergers and acquisitions do not seem to account for rising concentration nor do they confound the estimates of the effects of proprietary IT use.

Exposure to global trade might also confound the estimation (Melitz 2003, Autor et al. 2017). Column 3 includes a measure of industry import penetration ((imports–exports)/shipments) for NAICS manufacturing industries (Schott 2011) for 2002 through 2005. For non-manufacturing industries, I set import penetration to zero. This measure of import penetration has no effect on the coefficient of proprietary IT use and is not significantly correlated with industry concentration.

Columns 4 adds the average annual growth rate for real shipments from 1980 to 2002 for manufacturing industries.[23] It might be harder to maintain market share in a rapidly growing industry and rapidly growing industries might have greater need of IT. The coefficient on industry growth is negative and weakly significant (P = .077). The

[22] These are acquisitions by publicly listed firms of private and publicly listed firms. In aggregate, private firms do more acquisitions—85% of them in these data.

[23] Data from the NBER-CES Manufacturing Productivity database.

coefficient on proprietary IT use is larger, suggesting that, if anything, the omission of industry growth biases the coefficient downwards.[24]

Column 5 includes all of the right-hand side variable tested in columns 1-3 for the whole sample. The coefficient on the number of establishments in the industry is now statistically significant, but the coefficient on the software share remains roughly the same, suggesting that none of these additional variables confound the analysis of the role of IT.

Finally, I also included a measure of industry regulation in the regression (Al-Ubaydli and Mclaughlin 2015). I found no statistically significant relationship.

The Productivity Gap

The above data support the link between proprietary IT and industry concentration. If the paper's hypothesis is correct, proprietary IT should increase industry concentration by increasing the productivity gap between the top firms and the rest. The link between IT and a productivity gap should show up as a link between IT and labor productivity and also, in many industries such as retail, as a link between IT and establishment size.

Table 9 explores the relationship between the software share of the workforce and average establishment size, comparing the relationship for the top four firms in each industry with the relationship for the remaining firms. Because the Economic Census

[24] If I include the growth in the industry capital stock, that has a significant negative coefficient and the coefficient for the software share is even larger.

does not provide complete data for the manufacturing sector, that sector is necessarily excluded from the analysis that follows.

The table reports joint estimates using Zellner's "Seemingly Unrelated Regression" of equations relating the log of deflated revenues per establishment for each group of firms (Top 4 and the rest) separately:

$$\ln R_{it}^{top\ 4} = \alpha^{top\ 4} \cdot IT_{it} + \mu_i + \delta_t + \epsilon_{it}$$

$$\ln R_{it}^{rem} = \alpha^{rem} \cdot IT_{it} + \mu'_i + \delta'_t + \epsilon'_{it}$$

I use a log specification because establishment revenues are highly skewed. The first column shows the unrestricted regressions with controls for industry sector and year. The second column shows the regression where the coefficients for the industry sector and year dummies are constrained to be equal across equations. The bottom row reports a Wald test of the null hypothesis that $\alpha^{top\ 4} = \alpha^{rem}$.

In both columns, estimates of $\alpha^{top\ 4}$ and α^{rem} are both highly significant and the Wald test strongly rejects the null hypothesis. IT is strongly associated with greater revenue per establishment and the association is substantially stronger for the larger, presumably more productive, firms. These findings are consistent with the idea that IT brings scale economies to many industries.

Columns 3 and 4 repeat the analysis using log revenues per employee as the dependent variable. The results are broadly similar. Although this is not a causal analysis, these findings support the notion that IT may be implicated in the rising labor productivity gap between the top firms and the rest.[25]

[25] Note that revenues per employee includes the level of markups, so this is not a pure productivity measure.

Growth in Operating Margins

Some observers see rising profit margins (Figure 1) as evidence that competition has declined. How do these findings about rising concentration relate to the analysis of profits and markups? In theory, in long run equilibrium in a competitive market with homogenous productivity, firm operating margins should reflect only the returns needed to pay fixed capital costs. If margins were higher than that, new firms could profitably enter. Barkai (2016) presents evidence that firm margins have increased above and beyond payments to capital, concluding that this represents a decline in competition. These findings suggest some tension with the evidence found here regarding industry concentration.

However, if proprietary IT allows some firms to become more productive than others in the same industry, as above, then the more productive firms can earn quasi-rents. These would also be reflected in higher operating margins. Even in a competitive market, more productive firms could sell at the market price but profit from lower costs.

Some empirical analysis can help disentangle these effects. Table 10 provides an analysis of the growth in operating margins. The sample in this case consists of publicly listed US firms that reported in both 2000 and 2014, excluding firms in the finance sector.[26] The dependent variable is the change in operating margin between 2000 and 2014 where operating margin is defined as operating income after depreciation but before taxes, R&D, advertising and marketing expenditures all divided by revenues. I exclude R&D, advertising and marketing from income because I treat these as intangible

[26] In addition, the sample excludes the 5% tails in the dependent variable and firms where R&D spending exceeds 50% of revenues.

investments on the right-hand side of the regression equations. That is, operating profits should reflect the returns on investments in capital as well as returns to stocks of intangibles.

The operating margin for firm *i* at time *t* can be written

$$M_{it} = \alpha \cdot IT_{it} + \delta \cdot t + \beta_1 \frac{K_{it}^1}{R_{it}} + \beta_2 \frac{K_{it}^2}{R_{it}} + \cdots + \epsilon_{it}$$

where $K_i^1, K_i^2, \ldots$ represent stocks of capital assets as well as stocks of intangible assets, R&D and advertising and marketing. The β_j represent the rental rates for each type of capital. α represents the effect of IT. δ represents a time trend rate; if a general decline in competition were causing a rise in margins, then we should find $\delta > 0$. Because we are interested mainly in the growth of margins over this period (2000-2014) and because there are also likely significant firm fixed effects, I estimate the differenced equation over this interval:

$$\Delta M_i = \alpha \cdot \Delta IT_i + \delta + \beta_1 \Delta \frac{K_i^1}{R_i} + \beta_2 \Delta \frac{K_i^2}{R_i} + \cdots + \Delta\epsilon_i.$$

Table 10 reports some basic estimates. Column 1 reports a simple OLS regression and Column 2 reports the instrumental variable regression. Note that the IT measure is an industry-level measure while the other variables are for individual firms. In both columns the coefficients for software share are highly significant, but the IV estimate is substantially larger. At the sample mean for the change in software share (.007), these coefficients represent an increase in operating margins of 0.9% and 3.5% respectively. By comparison, the actual increase in operating margins for this sample is 3.2%, suggesting that IT can account for a major portion of the observed increase.

Column 3 repeats the regression in Column 2, but adds an additional variable, a measure of the change in industry regulation based on word counts in the Federal Code. If Federal regulation imposes substantial fixed compliance costs, then this might serve as an entry barrier, raising margins (Bessen 2016, Guttierez and Philippon 2017). There does seem to be a significant association between regulation and margins; at the sample mean, the increase in regulation may have contributed 1.6% to the growth in operating margins. But inclusion of this variable does not significantly alter the coefficient on IT share.

Finally, the constant term represents the background trend. This term is negative in all three specifications, significantly so in the third. It appears that once IT and intangibles are accounted for, the trend is not positive, contrary to the notion that a general decline in competition has led to rising firm margins. In any case, the evidence on operating margins does not seem to conflict with the findings above on industry concentration.

Conclusion

Firms are making large investments in proprietary information technology. The evidence in this paper suggests that these investments are changing industry structure and production. It is sometimes argued that information technology "levels the playing field" by providing inexpensive tools to small and young firms. This paper finds that much of the impact of IT may be, instead, to tilt the playing field in favor of those firms who are able to use it most effectively.[27] The use of proprietary IT is strongly associated with industry concentration across a wide range of sectors and the link is large enough to account for much of the recent rise in industry concentration. Instrumental variable regressions provide some support for the notion that this relationship is causal. This view is further supported by evidence that proprietary IT use is associated with greater labor productivity, especially among the top four firms in each industry. Proprietary IT is associated with a widening productivity gap between the top firms and the rest.

On the other hand, the observed increases in concentration are fairly modest. There are, of course, well known examples where IT facilitates highly concentrated markets as with Amazon's dominance in e-commerce. These cases may be "winner-take-all" markets. But the markets in this study show much lower levels of concentration and relatively small increases. While economies of scale or network effects might be at play in the markets studied here, it appears that there are limits to such scale effects; IT does not appear to generate a natural monopoly in most markets. These are "winner-take-a-bit-more" markets, consistent with the natural oligopoly models of Shaked and Sutton (1983, 1987). Perhaps more narrowly defined markets would be more likely to exhibit "winner-take-all" competition, but the market definitions used here from the Economic Census (at the 6-digit NAICS and higher level of aggregation) are the markets that have raised concern about growing concentration.[28]

---

[27] Some recent evidence suggests that cloud computing might be altering the relationship in favor of small firms (Jin and McElheran 2017).

[28] The Economist, "Too much of a good thing," March 26, 2016.

The findings of this paper suggest that much of the recent rise in industry concentration and much of the rise in firm operating margins can be attributed to the deployment of proprietary IT systems. A general decline in competition might also play a role in rising concentration and profits, but the evidence found here regarding competition is mixed. Merger and acquisition activity seems unrelated to industry concentration and the residual time trend in operating margins is not positive once intangible investments are taken into account. Overall, the analysis here suggests that the recent general rise in industry concentration is not mainly the result of anticompetitive activity that should worry antitrust authorities. While there may be other reasons to question antitrust policies (see, for instance, Kwoka 2012), the general rise in industry concentration does not appear to be a direct result of lax antitrust enforcement.

However, the effect of proprietary IT on industry structure does broach another concern: these changes in industry structure may dampen economic dynamism. For example, why aren't the productivity gains from IT shared more broadly beyond the top firms? Increasingly, it seems, top performing firms utilize new technologies productively while their rivals do not. Concentration appears to be rising because of "barriers to technology" if not actually barriers to entry. More research is needed to understand exactly how IT is related to the growing productivity gap. Top firms might be able to use patents and trade secrets to prevent the spread of new knowledge. Or perhaps, instead, top firms are better able to recruit and develop talented managers and workers skilled at working with the new systems. Or fixed costs might weaken incentives for laggard firms. Whatever the cause, the issue is important because the slow diffusion of new technologies might be related to sluggish aggregate productivity growth. Also, growing disparity in firm productivity might be related to growing inter-firm wage inequality. But the policies to address these issues, whether antitrust or other, depend very much on the diagnosis.

References

Akcigit, Ufuk, and Sina T. Ates. What Happened to US Business Dynamism?. No. w25756. National Bureau of Economic Research, 2019.

Al-Ubaydli, O. and McLaughlin, P.A., 2015. RegData: A numerical database on industry-specific regulations for all United States industries and federal regulations, 1997–2012. Regulation & Governance.

Andrews, Dan, Chiara Criscuolo, and Peter N. Gal. *The Best versus the Rest: The Global Productivity Slowdown, Divergence across Firms and the Role of Public Policy*. No. 5. OECD Publishing, 2016.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. The Fall of the Labor Share and the Rise of Superstar Firms. No. 23396. National Bureau of Economic Research, Inc, 2017.

Barkai, Simcha. "Declining labor and capital shares." Stigler Center for the Study of the Economy and the State New Working Paper Series 2 (2016).

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. "Industry Concentration in Europe and North America." (2019).

Basker, Emek. "The causes and consequences of Wal-Mart's growth." Journal of Economic Perspectives 21, no. 3 (2007): 177-198.

Berlingieri, Giuseppe, Patrick Blanchenay, and Chiara Criscuolo. "The great divergence." OECD working paper (2017).

Brynjolfsson, Erik, Andrew McAfee, Michael Sorell, and Feng Zhu. "Scale without mass: business process replication and industry dynamics." (2008).

Calligaris, Sara, Chiara Criscuolo, and Luca Marcolin. Mark-ups in the digital era. No. 2018/10. OECD Publishing, 2018.

Crouzet, Nicolas, and Janice Eberly. "Intangibles, Investment, and Efficiency." In AEA Papers and Proceedings, vol. 108, pp. 426-31. 2018.

De Loecker, Jan, and Jan Eeckhout. The Rise of Market Power and the Macroeconomic Implications. No. w23687. National Bureau of Economic Research, 2017.

De Loecker, J. and Eeckhout, J., 2018. Global market power (No. w24768). National Bureau of Economic Research.

Decker, Ryan A., John C. Haltiwanger, Ron S. Jarmin, and Javier Miranda. Changing business dynamism and productivity: Shocks vs. responsiveness. No. w24236. National Bureau of Economic Research, 2018.

Demsetz, Harold. "Industry structure, market rivalry, and public policy." The Journal of Law and Economics 16, no. 1 (1973): 1-9.

Diez, Mr Federico, Mr Daniel Leigh, and Suchanan Tambunlertchai. Global Market Power and its Macroeconomic Implications. International Monetary Fund, 2018.

DiNardo, John E., and Jörn-Steffen Pischke. "The returns to computer use revisited: Have pencils changed the wage structure too?." The Quarterly Journal of Economics 112, no. 1 (1997): 291-303.

Dunne, Timothy, Lucia Foster, John Haltiwanger, and Kenneth R. Troske. "Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment." Journal of Labor Economics 22.2 (2004): 397-429.

Eckard, E. Woodrow. "Plant-level scale economies and industrial concentration." The Quarterly Review of Economics and Finance 34, no. 2 (1994): 173-182.

Ellickson, Paul B. "Quality competition in retailing: A structural analysis." International Journal of Industrial Organization 24, no. 3 (2006): 521-540.

England, Paula, and Barbara Kilbourne. Occupational Measures from the Dictionary of Occupational Titles for 1980 Census Detailed Occupations. Ann Arbor, MI: Inter-university Consortium for Political and Social Research, 2013.

Grullon, Gustavo, Yelena Larkin and Roni Michaely. "Are US Industries Becoming More Concentrated?" working paper (2017).

Gutiérrez, Germán and Thomas Philippon. "Declining Competition and Investment in the U.S." NBER Working Paper (2017).

Gutiérrez, Germán, and Thomas Philippon. How EU Markets Became More Competitive Than US Markets: A Study of Institutional Drift. No. w24700. National Bureau of Economic Research, 2018.

Hall, B.H., Berndt, E. and Levin, R.C., 1990. The Impact of Corporate Restructuring on Industrial Research and Development. Brookings Papers on Economic Activity. Microeconomics, 1990, pp.85-135.

Harrigan, James, Ariell Reshef, and Farid Toubal. *The march of the techies: Technology, trade, and job polarization in France, 1994-2007*. No. w22110. National Bureau of Economic Research, 2016.

Jäger, Kirsten. "EU KLEMS Growth and Productivity Accounts 2017 Release," Description of Methodology and General Notes (2017).

Jin, Wang, and Kristina McElheran. "Economies before Scale: Productivity and Survival of Young Plants in the Age of Cloud Computing." 2017.

Lewellen, W.G. and Badrinath, S.G., 1997. On the measurement of Tobin's q. *Journal Of Financial Economics*, *44*(1), pp.77-122.

Melitz, Marc J. "The impact of trade on intra‐industry reallocations and aggregate industry productivity." Econometrica 71, no. 6 (2003): 1695-1725.

Peltzman, Sam. "Productivity and Prices in Manufacturing During an Era of Rising Concentration," working paper (2018).

Rinz, Kevin. "Labor market concentration, earnings inequality, and earnings mobility." Center for Administrative Records Research and Applications Working Paper 10 (2018).

Rognlie, Matthew. "Deciphering the fall and rise in the net capital share: accumulation or scarcity?." Brookings papers on economic activity 2015.1 (2016): 1-69.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. Diverging trends in national and local concentration. No. w25066. National Bureau of Economic Research, 2018.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

Schott, Peter K. "U.S. Manufacturing Exports and Imports by SIC or NAICS Category and Partner Country, 1972 to 2005," (2011) http://faculty.som.yale.edu/peterschott/sub_international.htm.

Shaked, Avner, and John Sutton. "Natural oligopolies." Econometrica: journal of the Econometric Society (1983): 1469-1483.

Shaked, Avner, and John Sutton. "Product differentiation and industrial structure." The Journal of Industrial Economics (1987): 131-146.

Shapiro, Carl. "Antitrust in a Time of Populism," working paper (2017).

Tambe, Prasanna, and Lorin M. Hitt. "The productivity of information technology investments: New evidence from IT labor data." *Information Systems Research* 23.3-part-1 (2012): 599-617.

Unger, Gabriel. ""Scale-Biased Technological Change" working paper (2019).

U.S. Department of Labor. *Dictionary of Occupational Titles,* fourth edition. Government Printing Office, 1977.

Villalonga, B., 2004. Intangible resources, Tobin's q, and sustainability of performance differences. *Journal of Economic Behavior & Organization*,*54*(2), pp.205-230.

White, Lawrence and Jasper Yang. 2017. "What Has Been Happening to Aggregate Concentration in the U.S. Economy in the 21st Century?" working paper.

Tables

Table 1. Summary Statistics

IT occupations, share of hours worked	2.2%	
Percent of industries where top 4 firms > 50% of revenues	15.1%	
Share of industry revenue going to:		
Top 4 firms	27.8%	
Top 8 firms	36.0%	
Top 20 firms	46.6%	
Top 50 firms	55.9%	
Average five-year change, 1997-2007:		
Change in share of industry revenue going to:		
Top 4 firms	1.43%	
Top 8 firms	1.60%	
Top 20 firms	1.67%	
Top 50 firms	1.70%	
Median Characteristics (excludes manufacturing)	Industry	Top 4 firms
Revenues / establishment (1000s $2009)	$1,706.6	$7,247.9
Revenues / employee (1000s $2009)	$146.4	$194.8
Average annual pay (1000s $2009)	$32.3	$36.7
Wage bill / revenues	23.5%	19.4%

Note: Sample for levels includes 808 observations of industries with IT share data over the years 1997, 2002, 2007, and 2012; sample for changes in concentration ratios is 335; sample for industry characteristics excludes manufacturing because Economic Census does not report number of establishments for top 4 firms. Dollar figures are deflated by the GDP Deflator for 2009 = 1.

Table 2. First stage regressions

Dependent variable: IT share of workforce			
Year	2002	2007	2012
Sedentariness	7.03** (2.08)	5.14** (1.26)	5.12** (1.17)
Sector dummies	✓	✓	✓
No. of observations	97	108	100
Adjusted R-squared	0.142	0.210	0.296
Simple correlation	.312	.307	.328

Note: Standard errors in parentheses, * = significant at 5% level; ** = significant at 1% level. Sectors are 1-digit NAICS sectors. The samples from 2002 through 2012 are only those industries that have both a sedentariness index from the 1980 Census and IT share of the workforce from the ACS.

Table 3. Do industry worker characteristics affect concentration ratios?

Dependent variable	Share of revenues, 2002, 2007, 2012				Change in revenue share, 1977-2002
	Top 4 firms	Top 8 firms	Top 20 firms	Top 50 firms	Top 4 firms, manufacturing
Industry characteristics:					
Share professional & managers	0.07 (0.18)	0.09 (0.21)	0.12 (0.24)	0.07 (0.25)	0.10 (0.44)
Mean years school	-2.44 (3.87)	-3.04 (4.79)	-4.22 (5.04)	-4.94 (4.82)	9.99 (11.79)
Log wage	24.95 (14.89)	29.27 (16.83)	34.31 (16.65)°	32.21 (14.87)°	-32.49 (43.97)
Industry digit dummies	✓	✓	✓	✓	
Year dummies	✓	✓	✓	✓	
Sector dummies	✓	✓	✓	✓	
Observations	669	669	671	666	69
R-squared	0.241	0.259	0.303	0.324	0.071
Joint test (P value)	0.455	0.441	0.303	0.157	0.170

Note: Robust standard errors in parentheses, ° = significant at 10% level; * = significant at 5% level; ** = significant at 1% level. Details of the variables and samples described below.

Table 4. Placebo tests

Sample	Manufacturing industries		Compustat firms	
Dependent variable	4-firm concentration ratio		Operating margin	
	1977	1997-2012	1977	2000-2014
Sedentariness	0.19 (0.21)	1.06 (0.20)**	0.07 (0.05)	0.27** (0.02)
Year dummies		✓		✓
SIC2 dummies			✓	✓
Capital and intangible stocks			✓	✓
Observations	79	185	1179	31346
R-squared	0.012	0.200	0.651	0.625

Note: Robust standard errors in parentheses, * = significant at 5% level; ** = significant at 1% level. The sedentariness index is assigned to firms via the Census NAICS classification; consequently, the firms in the 1977 sample also appear in year 1998 or later when NAICS codes were assigned. As in the analysis below, the 1% tails of the dependent variable were excluded. The firm regression is weighted by real sales, also corresponding to the analysis below.

Table 5. Regressions on Concentration Ratios

Dependent Variable: Concentration Ratio, 2002 – 2012				
A. OLS	Top 4 firms	Top 8 firms	Top 20 firms	Top 50 firms
Software share	2.14** (0.32)	2.66** (0.36)	3.12** (0.40)	2.88** (0.44)
No. of observations	725	725	727	722
R-squared	.184	.175	.155	.104
B. Sector Fixed Effects				
Software share	1.99** (0.31)	2.39** (0.34)	2.71** (0.36)	2.40** (0.38)
No. of observations	725	725	727	722
R-squared	.257	.281	.326	.337
Within R-squared	.179	.168	.150	.103
C. Sector Fixed Effects, weighted by shipments/revenues				
Software share	1.27** (0.35)	2.02** (0.40)	3.08** (0.42)	3.88** (0.42)
No. of observations	720	720	722	717
R-squared	.222	.256	.305	.347
Within R-squared	.128	.148	.167	.176
Dependent Variable: Concentration Ratio, 1997 – 2012				
D. Reduced form IV				
Sedentariness	19.81** (7.48)	33.63** (8.74)	53.70** (10.29)	60.55** (11.25)
No. of observations	1829	1842	1838	1816
R-squared	0.163	0.190	0.222	0.243
Scaled coefficient, SW share	3.17	5.39	8.60	9.70
Dependent Variable: Five-year Change in Concentration Ratio, 1997 – 2007				
E. Reduced form IV	Top 4 firms	Top 8 firms	Top 20 firms	Top 50 firms
Sedentariness	3.49+ (2.07)	3.70* (1.85)	3.70* (1.58)	1.89 (1.18)
No. of observations	844	854	846	832
Scaled coefficient, SW share	0.56	0.59	0.59	0.30

Note: Standard errors in parentheses, + = significant at 10% level, * = significant at 5% level; ** = significant at 1% level. Standard errors are heteroskedastic-robust. All level regressions include year dummies and industry digit dummies. Fixed effects regressions include 1-digit NAICS controls. Dependent variable is share of revenues accounted for by top firms (varying number). The reduced form IV regresses the dependent variable on the instrument, a measure of the sedentariness of the industry workforce, using occupational measures from 1977 apportioned to industries using the 1980 Census. The IV regressions are weighted by shipments/revenues. The instrument is not available for the same sample of industries as used in the OLS estimates. The scaled coefficient is determined by dividing the sedentariness coefficient by the coefficient obtained by regressing the software share of the workforce on sedentariness with year and sector fixed effects for those industries that have both measures. The scaling factor is 6.24.

Table 6. Alternative Instruments

Dependent Variable: Concentration Ratio, 1997 – 2012				
A. Reduced form IV				
1980 SW share of workforce	6.34** (1.56)	9.95** (1.73)	14.65** (2.03)	16.52** (2.24)
No. of observations	1829	1842	1838	1816
R-squared	0.181	0.223	0.279	0.311
Scaled coefficient, SW share	2.75	4.32	6.36	7.17
B. Two-stage Least Squares				
EU SW share of investment	0.66* (0.28)	0.75* (0.34)	0.87* (0.39)	1.00* (0.43)
No. of observations	72	72	72	72
R-squared	0.004	0.008	0.014	0.010

Note: Standard errors in parentheses, * = significant at 5% level; ** = significant at 1% level. Standard errors are heteroskedastic-robust and regressions are weighted by shipments/revenues. Panel A includes year dummies, sector dummies, and industry digit dummies; Panel B includes year dummies. Panel A regresses industry concentration against the software share of the workforce obtained from the 1980 Census, using the walkway to NAICS industries. The instrument is not available for the same sample of industries as used in the OLS estimates in Table 5. The scaled coefficient is determined by dividing the regression coefficient by the coefficient obtained by regressing the software share of the workforce on sedentariness with year and sector fixed effects for those industries that have both measures. The scaling factor is 2.30. Panel B uses the share of software in investment for 18 EU countries as an instrument in two-stage least squares.

Table 7. Long Difference in Four-firm Concentration Ratio

Dependent Variable: Change in Four Firm Concentration Ratio			
	Manufacturing only, 1977 - 2002		
	OLS	OLS	IV
Software share, 1980	8.98 (1.43)**		7.59 (3.09)*
Change in software share		1.76 (1.05)°	
No. of observations	71	71	71
R-squared	0.154	0.053	0.15
Mean IT variable	0.55	1.55	0.55
Average effect	4.90	2.74	4.14

Note: Standard errors in parentheses, ° = significant at 10% level; * = significant at 5% level; ** = significant at 1% level. OLS errors are robust to heteroscedasticity; IV errors are bootstrapped. Dependent variable is the change in share of revenues accounted for by top 4 firms. Software share is instrumented using a measure of the sedentariness of the industry workforce, using occupational measures from 1977 apportioned to industries using the 1980 Census. Excludes the 1% tails of the dependent variable. The null hypothesis that software share is exogenous in the IV regression cannot be rejected (P = .352)

Table 8. Possibly Confounding Variables

Dependent Variable: Four Firm Concentration Ratio					
	1	2	3	4	5
Software share	1.88** (0.31)	1.80** (0.35)	2.00** (0.32)	3.15** (0.64)	1.70** (0.33)
Number of establishments (1000s)	-0.00 (0.00)*				-0.00 (0.00)**
M&A index, 1985-2001		-2.79 (2.78)			-3.01 (2.53)
Import penetration			-1.85 (3.99)		-1.34 (3.76)
Output growth, 1980-2002				0.05 (1.06)	
Industry digit dummies	✓	✓	✓	✓	✓
Year dummies	✓	✓	✓	✓	✓
Sector dummies	✓	✓	✓	✓	✓
No. of observations	724	661	725	276	660
R-squared	0.287	0.274	0.257	0.401	.302

Note: Robust standard errors, in parentheses, ° = significant at 10% level; * = significant at 5% level; ** = significant at 1% level. OLS regressions on pooled industries for 2002, 2007, 2012.

Table 9. Establishment size, labor productivity, and IT

Dependent variable:	Log Revenues / establishment ($2009)		Log Revenues / employee ($2009)	
	Unrestricted	Restricted	Unrestricted	Restricted
Top 4 firms				
Software share	0.25 (0.03)**	0.48 (0.03)**	0.15 (0.02)**	0.25 (0.02)**
Year dummies	✓	✓	✓	✓
Sector dummies	✓	✓	✓	✓
No. of observations	439	439	439	439
R-squared	0.256	0.025	0.296	0.212
Remaining firms				
Software share	0.14 (0.02)**	0.07 (0.03)**	0.11 (0.02)**	0.13 (0.02)**
Year dummies	✓	✓	✓	✓
Sector dummies	✓	✓	✓	✓
No. of observations	439	439	439	439
R-squared	0.292	0.245	0.359	0.353
Test equality of SW share coefficients (Prob. value)	0.000	0.000	0.001	0.000

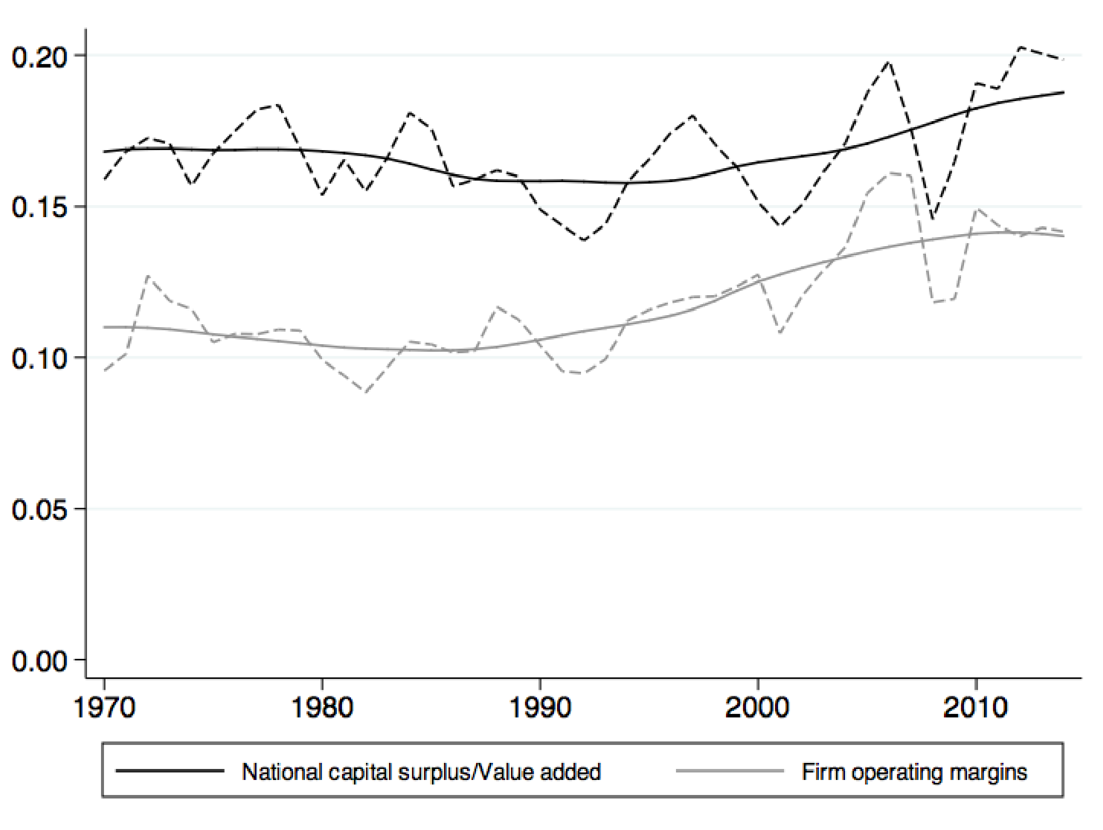
Note: **=significant at 1% level; *=significant at 5% level. Estimates use the Seemingly Unrelated Regression model for separate equations for the top 4 firms in each industry and for the remaining firms in each industry. The sample excludes manufacturing industries (data was not reported). The restricted estimates constrain the coefficients of the dummy variables to be equal across the two equations. The bottom row reports the probability of the null hypothesis in a Wald test that the coefficients of software share are equal across the two equations.

Table 10. Change in Operating Margins, 2000 – 2014

Dependent Variable: Δ Operating income after depreciation before taxes, R&D, advert. / Revenues						
	OLS		IV		IV	
Δ IT share	2.71*	(1.34)	6.71	(5.81)	6.59**	(2.38)
Δ Capital stock	0.01**	0.00	0.00	(0.01)	0.00	(0.00)
Δ R&D stock	0.06*	(0.02)	-0.04**	(0.00)	-0.04**	(0.00)
Δ Advertising stock	0.47**	(0.05)	0.51**	(0.08)	0.52**	(0.04)
Δ Regulation					0.07*	(0.03)
Constant	-0.01	(0.01)	-0.02	(0.01)	-0.03**	(0.01)
No. observations	912		1000		840	
R-Squared	.255		.188		.207	

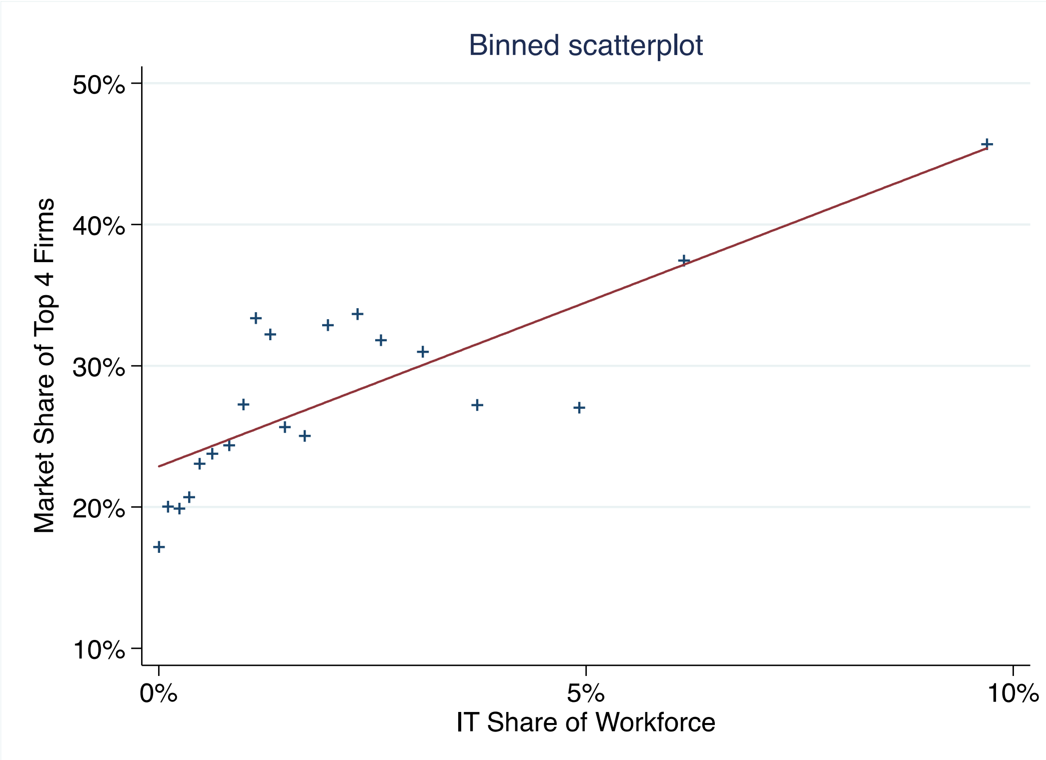
Note: **=significant at 1% level; *=significant at 5% level. Standard errors are clustered by industry. Sample is all US Compustat firms excluding 5% tails of the dependent variable and firms where R&D > .5*sales. IV uses sedentariness index as instrument.

Figure 1. Operating Margins



Note: Solid lines are kernel smoothed. Black line is from the System of National Accounts, Bureau of Economic Analysis. It shows the ratio of the net operating surplus to gross value added for the corporate sector (nonfinancial and financial). The gray line is the ratio of aggregate operating income after depreciation before taxes to revenues for firms publicly listed in the US.

Figure 2. 4-Firm Concentration Ratio and IT Share of Workforce



Note: For 254 industries excluding IT-producing industries over years 1997, 2002, 2007, and 2012.

Appendix

Table A1. Distribution of observations across sectors

Sector	Percent of sample	Change in four-firm concentration ratio, 2002-2007	Software share of workforce
Mining, utilities, construction	1.6	0.00	2.4%
Manufacturing	38.6	0.17	2.4%
Wholesale, retail, transportation, warehousing	25.9	2.23	1.6%
Finance, real estate, business services	17.0	1.84	3.7%
Education, health	8.6	-0.77	1.3%
Recreation, hotel, food services	3.7	1.13	0.6%
Other services	4.5	-0.15	0.9%

Table A2. Sedentariness across sectors

Sector	Lowest and Highest Industry	Sedentariness	Correlation with IT share
Agriculture		0.14	--
	Animal production	0.04	
	Forestry, except logging	0.39	
Mining, utilities, construction		0.36	0.847
	Coal mining	0.12	
	Not specified utilities	0.58	
Manufacturing		0.30	0.876
	Animal slaughtering and processing	0.12	
	Aerospace products and parts	0.73	
Wholesale, retail, transportation, warehousing		0.51	0.245
	Pipeline transportation	0.13	
	Jewelry, luggage, and leather goods stores	0.94	
Finance, real estate, business services		0.70	0.663
	Other administrative, and other support services	0.08	
	Architectural, engineering, and related services	0.98	
Education, health		0.49	0.164
	Child care	0.09	
	Office of chiropractors	0.96	
Recreation, hotel, food service		0.30	0.761
	Drinking places, alcohol beverages	0.06	
	Independent artists	0.74	
Other services		0.36	0.007
	Beauty salons	0.08	
	Nail salons and other personal care services	0.92	

Table A3. Four-firm concentration ratio by industry level

	3 digit	4 digit	5 digit	6 digit
Software share	2.28 (1.24)°	0.54 (0.33)°	2.40 (0.99)*	6.30 (0.98)**
Year dummies				
No. of observations	75	458	150	45
R-squared	0.046	0.006	0.047	0.679

Note: Robust standard errors in parentheses, ° = significant at 10% level; * = significant at 5% level; ** = significant at 1% level.