Appendix

A Appendix for Section 2: Measurement

This section provides additional results related to the mark-up estimates of De-Loecker et al. (2019) (DLEU hereafter). We begin with a brief discussion of the accounting definition of COGS, and its implications for mark-up estimation; followed by a discussion of technological change and it's relation to Sales, General and Administrative (SG&A) expenditures.

A.1 Accounting Definitions

Under the methodology of De Loecker and Warzynski (2012), mark-up estimates are unbiased as long as the variable input used in the estimation is indeed variable, and is consistently defined over time. Finding such a measure is not trivial, particularly in accounting statements. DLEU use COGS as their variable input which, according to GAAP, is defined as "the cost of inventory items sold during a given period." This is clearly defined for businesses that make, buy or sell goods to produce income, such as manufacturing, retail and wholesale trade. It is much less clear for service and information businesses. Pure service companies such as accounting firms, law offices, business consultants and many information technology firms have no goods to sell and therefore no inventory. As a result, they do not even report COGS on their income statement.[27] Some of them report only more granular line items, while others report "Cost of Revenues" instead. Importantly, cost of revenues includes the cost of delivering a product or service in addition to producing it, hence is broader than COGS. Such ambiguity in accounting definitions, coupled with changes in the nature of production, gives firms discretion on what is included in COGS vs. SG&A. Ultimately, this leads to the inclusion of some (quasi-)fixed expenditures in COGS, as well as changes in the definition of COGS over time – both of which may violate the assumptions underlying DLEU. Two examples:

Consider Delta Airlines, which does not report COGS in it's annual statements. Instead, Compustat creates a measure of COGS by combining a series of granular line items. Such items include clearly variable expenses such as aircraft fuel and landing fees – but also quasi-fixed expenses such as aircraft rent expense (typically associated with long term leases) and head-office salaries and profit sharing expenses (typically included in SG&A).

Google (Alphabet Inc), on the other hand, reports Cost of Revenues. The largest component of Cost of Revenues are traffic acquisition costs (TAC), which are identifiable, direct costs attributable to production. They roughly match the definition of COGS. However, Cost of Revenues also includes "expenses associated with our data centers and other operations (including bandwidth, compensation expense (including stock-based-compensation), depreciation, energy, and other equipment costs)." Clearly, data center and operation expenditures include long term investment in tangible and intangible assets indirectly related to the delivery of services (e.g., software, organizational capabilities, equipment). Again, this may violate the variable cost assumption underlying DLEU. Moreover, Google can exercise discretion on what is classified as SG&A

[27] See link for example, which lists personal service businesses that do not report COGS.

Table 7: *Summary of Income Statement (as % of sales)*

	1970-1979	1980-1989	1990-1999	2000-2009	2010-2017	$\Delta 00s - 70s$
Sale[a+b+c]	100	100	100	100	100	NA
COGS[a]	72.5	70.9	66.1	66.3	65.6	**-7.0**
SG&A[b]	14.4	16.4	19.3	18.4	17.9	**3.5**
OIBDP[c]	13.1	12.8	14.7	15.4	16.6	**3.5**
DP[d]	3.6	4.4	4.8	4.8	4.9	**1.3**
OIADP[c-d]	9.5	8.4	9.9	10.6	11.8	**2.2**
Intan K share (BEA)	6.2	7.5	11.1	12.5	13.4	**7.2**
Intan K share (PT)	28.2	33.3	38.5	47.3	49.0	**20.8**
Firm x year pairs	41045	49809	65295	55549	33304	

Notes: Annual data. Table shows the weighted average share of each income statement line item as a percent of sales. Source: Compustat for a, b, c and d. BEA and Peters and Taylor 2016 for the share of Intangible Capital.

instead of Cost of Revenues. In fact, Google reported stock-based-compensation separate from Cost of Revenues up to 2005 but combined it after 2006.

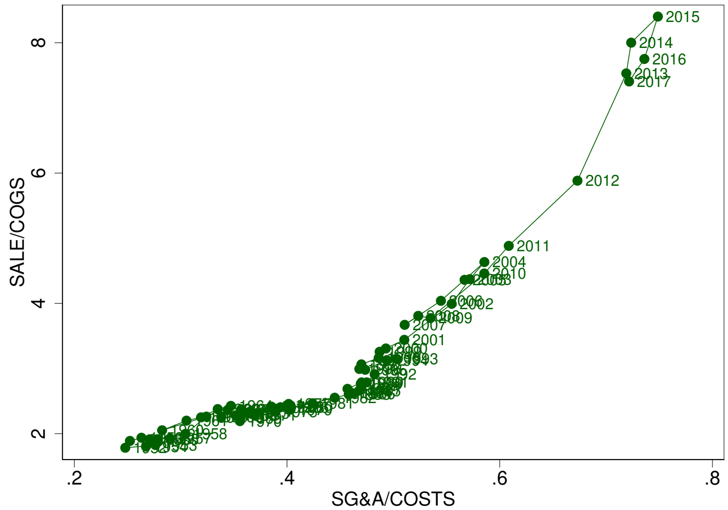
A.2 Role of SG&A and Intangibles

The above issues related to the measurement of variable costs – as well as the treatment of SG&A – pose significant challenges for the estimation. However, even assuming that COGS is a perfect proxy of variable costs and that SG&A is properly accounted for in the production function estimation, there is a more fundamental issue with the interpretation of mark-ups as a proxy of market power: technological change and the rise of fixed costs.

The share of SG&A in total costs has increased over the past 30 years, precisely when the share of COGS has fallen. Table 7 summarizes this fact, by showing the weighted average share of key income statement line items as a percent of sales. The COGS-share of sales declined by nearly 7 percentage points, while the SG&A and depreciation shares increased by 3.5 and 1.3 percentage points, respectively. Thus, most of the decrease in COGS was offset by a rise in SG&A and DP. But operating profits after depreciation also increased, by 2.2 percentage points of sales. The increase in SG&A and depreciation are consistent with a shift towards intangible capital: SG&A includes most intangible-building activities such as R&D, Advertising and Software-development expenses; and intangibles have higher depreciation rates (Corrado and Hulten, 2010). Most SG&A expenses are fixed in the short-run, which requires a careful treatment while estimating production functions. This is the subject of an ongoing debate (Traina, 2018; Karabarbounis and Neiman, 2018).

To understand the significance of rising SG&A for mark-up estimation, figure 16 shows the sales-weighted average of SALE/COGS and SG&A cost-shares (SG&A/COSTS) for firms in the the top quantile of the SALE/COGS distribution each year. As shown, SALE/COGS increased precisely at the firms where the SG&A cost-share increased – which points towards a major technological change, likely involving a rise in fixed costs. This has significant implications for the interpretation of mark-ups as a measure of market power. Two examples.

Figure 16: *Average SALE/COGS vs. SG&A intensity for high SALE/COGS firms*



Notes: scatter plot of the weighted average SALE/COGS and SG&A cost-share across all Compustat firms in the top quantile of the SALE/COGS distribution, by year.

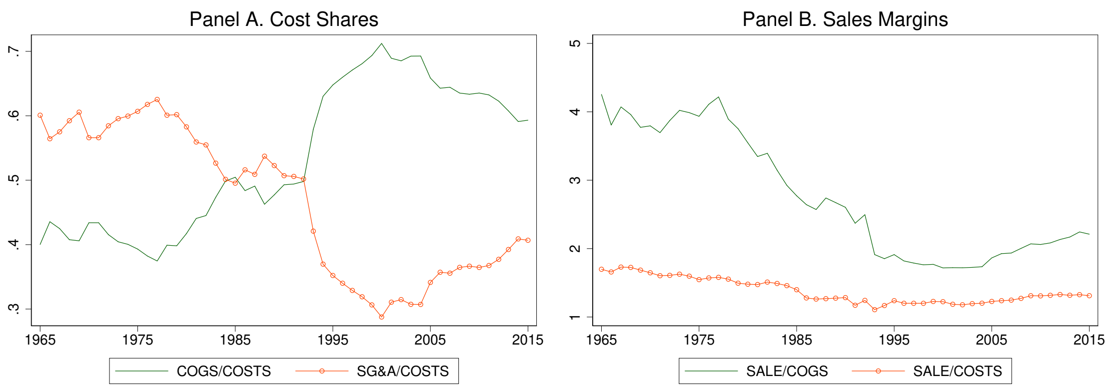
IBM. Consider IBM, a firm that transitioned from providing mostly products to mostly services, beginning in 1994. As shown in Panel A of Figure 17, the cost-share of COGS increased from 40 to 60% while the cost-share of SG&A decreased by a similar amount, precisely as IBM transitioned from a high-overhead, low-COGS business model (Hardware) to a high-COGS, low overhead business model (Consulting, where staff expenditures are included in COGS).[28] The implied mark-up fell sharply from 4 to 2 (Panel B). Does this mean that the extent of competition faced by IBM increased sharply from 1965 to 2015? Probably not. In the long-run, IBM's ratio of SALE to COGS is dominated by it's SG&A intensity, which is in turn dictated by its product mix. It tells us much about IBM's production function and it's share of fixed vs. variable costs, but less about the extent of (dynamic) competition faced by IBM in product markets. In fact, while IBM's SALE/COGS ratio fell by 48% from 1965 to 2015, margins (SALE/COSTS) fell by only 10%.

Walmart. IBM is interesting because the firm transitioned across widely different business models (curiously in the opposite direction of the economy, from a high SG&A to a high COGS model). A very different example is Walmart: a firm that maintained it's business model but invested heavily in intangible assets to improve logistics and gain market share (Panel A of Figure 18). This is consistent with IT investments driving concentration, as described in Bessen (2017). SALE/COGS increased rapidly with SG&A, yet profit margins (and the relative price of retail trade) actually fell.

These are specific examples, but as shown in Figure 6 above, the divergence between SALE/COGS and

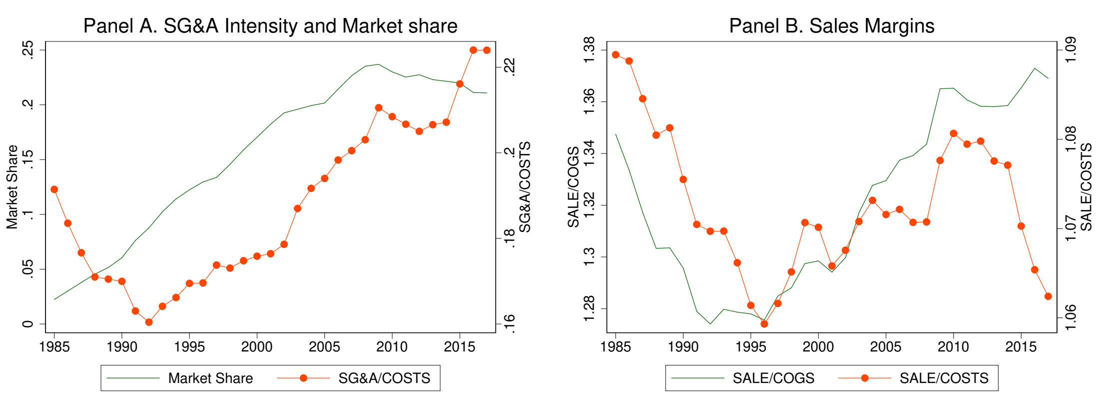
[28] The composition of COGS also changed, likely affecting the elasticity of sales to COGS. In 1992, costs associated with hardware and software sales accounted for 36.9% of sales. By 2016, the same figure dropped to only 8.2% of sales. Costs associated with services increased from 9.4% to 42.6%. IBM was eventually re-categorized from NAICS 3341 (Computer and peripheral equipment manufacturing) to 5415 (Computer Systems Design and Related Services) in 1998 and to 5191 (Other information services) in 2016. It is not clear to us how the change in industry categorization is dealt with by DLEU, but neither using a constant elasticity nor changing IBM from one industry to another in a particular year is entirely satisfactory – though this is a standard problem whenever industry segments are used.

Figure 17: *IBM: Cost Shares and Sales Margins*



Source: Compustat NA. COSTS = COGS + SG&A.

Figure 18: *Walmart: Cost Shares, Market Shares and Sales Margins*



Source: Compustat NA. Market share for BEA Retail Trade industry.

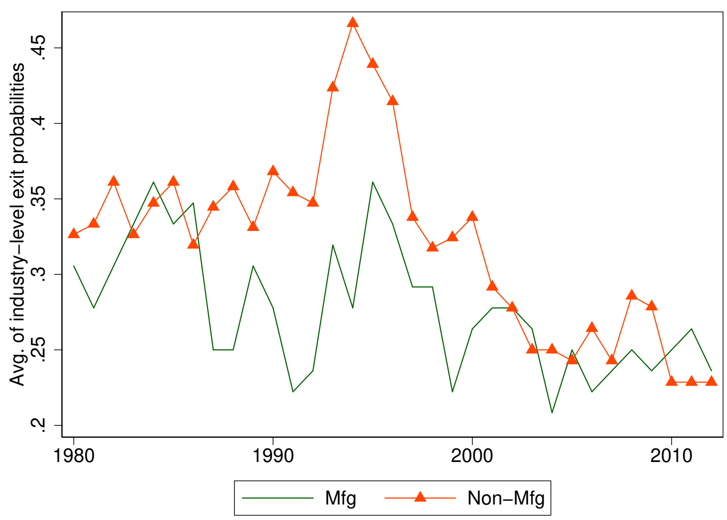
profits remains at the country-level. As a result, rising COGS-based mark-ups – by themselves – tell us little about the long-run evolution of competition and market power. DLEU acknowledge as much, noting that*"technological change will lead to higher mark-ups (due to lower marginal costs), but prices will not drop because firms need to generate revenue to cover fixed costs. As a result, profits will continue to be low and* ***higher mark-ups do not imply higher market power.****"* Profits – therefore – remain the only reliable measure of marker power; and the one we focus on here and in related work.

B Appendix for Section 3: Aggregate Evidence

B.1 Entry, Exit and Turnover

Figures 19 replicates figure 7 using market value and separating manufacturing and non-manufacturing industries. As shown, the drop in turnover is more pronounced for non-manufacturing industries.

Figure 19: *MV-based Leader Turnover, by Sector*



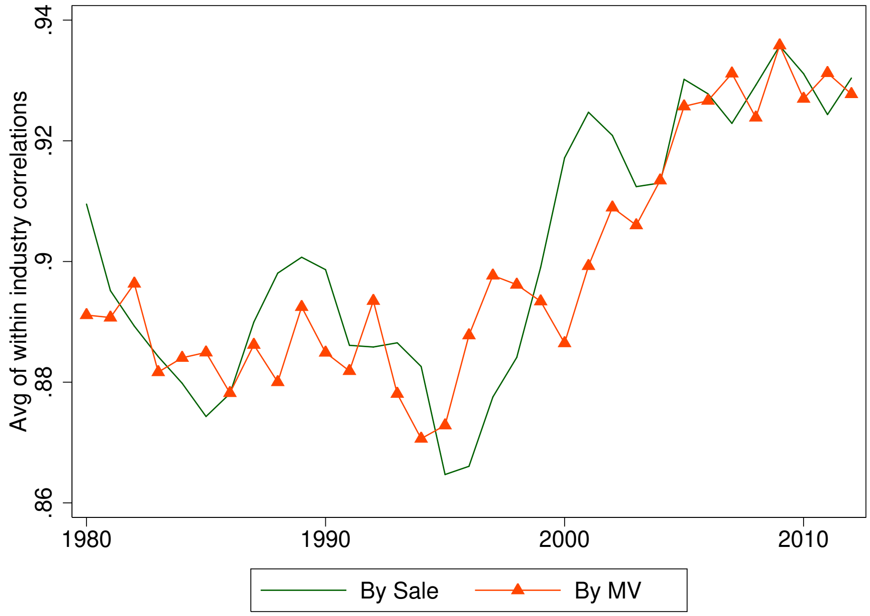
Source: Compustat NA, following BEA industries. Includes only industry-years with 5 or more firms. See text for details.

Figure 20 presents an additional measure of turnover, based on the correlation of firm rankings over time. For a particular measure Z (sales, market value, etc.), we define

$$RkCorr = Corr_{i\in j}\left(rank\left(z_{i,j,t}\right); rank\left(z_{i,j,t+5}\right)\right),$$

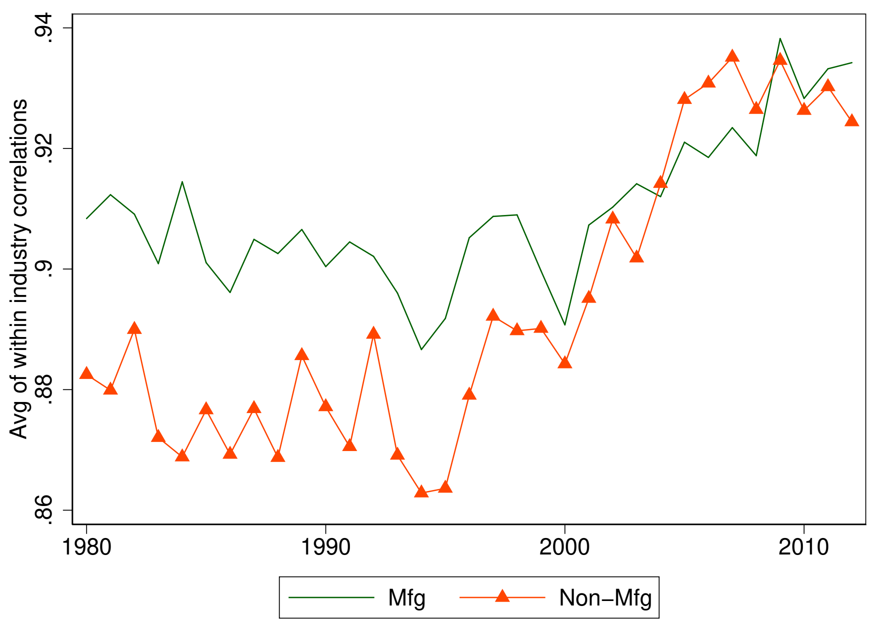
where $rank(z_{i,j,t})$ is the rank of firm i in industry j at time t according to the measure z. We again find a sharp increase in persistence after 2000. Figure 21 presents the same results but separating manufacturing and non-manufacturing sectors.

Figure 20: *Correlation of 5Y-ahead Firm Ranks*



Source: Compustat NA, following BEA industries.Only industry-years with 5 or more firms are included. See text for details.

Figure 21: *MV-based correlation of 5Y-ahead rankings by sector*

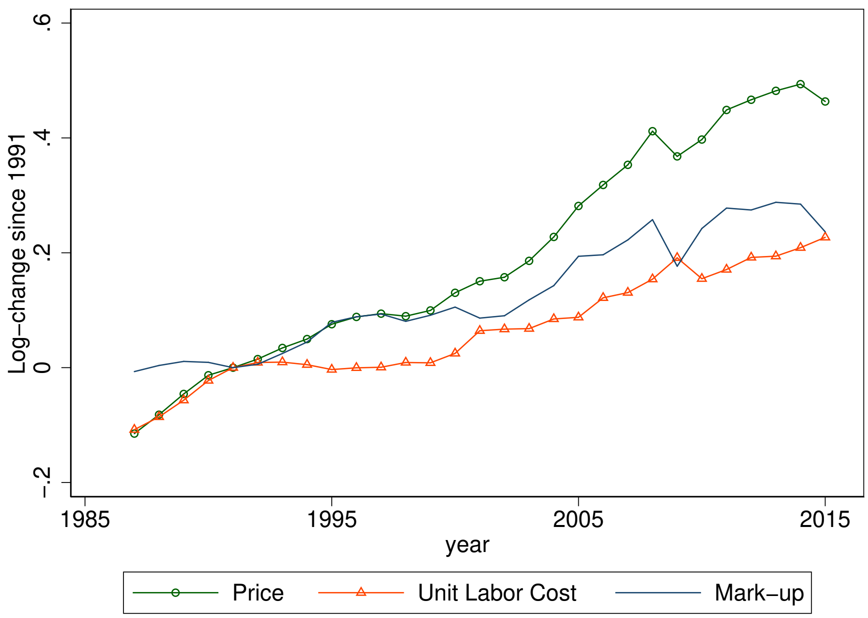


Source: Compustat NA, following BEA industries. Includes only industry-years with 5 or more firms. See text for details.

B.2 Concentration, Productivity and Prices

We are interested in decomposing the correlation between concentration and mark-ups into the underlying components: prices, wages and labor productivity. In Figure 22 we plot the aggregate evolution of prices and unit labor costs since 1989. As shown, prices increased faster than unit labor costs, leading to an increase in mark-ups.

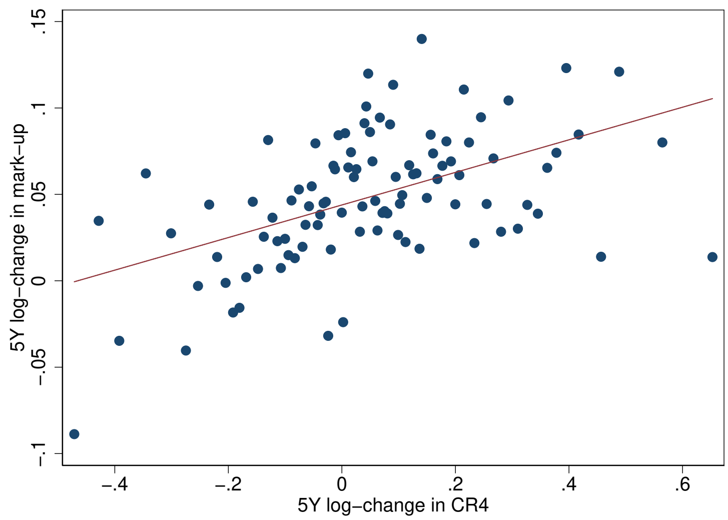
Figure 22: *Prices, ULC and Mark-ups in US*



Notes: weighted average change in prices, per-unit labor costs and mark-ups (computed as the residual) across all industries in our sample. Based on BLS multifactor tables.

Figure 23 provides a bin-scatter plot of changes in mark-ups against changes in CR4. As shown, the relationship is quite robust.

Figure 23: *Mark-ups vs. Concentration*



Notes: Concentration from US Economic Census. Mark-ups from the NBER-CES database for manufacturing and the Economic Census (output, employment and wages) and the BEA detailed GDP By Industry Accounts (prices). See Section 3.2 for details.

Last, Table 8 reports regressions of the following form using our detailed industry dataset of prices and productivity:

$$\Delta_5 log(Y_{jt}) = \beta \Delta_5 log(CR4) + \gamma_{s,t} + \varepsilon_{jt}.$$

where j denotes industries and t denotes years. $\gamma_{s,t}$ denotes sector x year fixed effects. To facilitate comparison to Ganapati 2018, we standardize $\Delta_5 log(CR4)$ to have mean zero and variance one. Outcome variables Y_{jt} are based on the following interlinked outcomes:

$$\begin{aligned} \Delta_5 \log \mu &= \Delta_5 \log P - \Delta_5 \log ULC, \\ &= \Delta_5 \log P - [\Delta_5 \log w - \Delta_5 \log LP] . \end{aligned} \tag{6}$$

Panel A includes all industries, while Panels B and C separate manufacturing and non-manufacturing industries. In line with Autor et al. 2017b and Ganapati 2018, concentration is positively correlated with labor productivity growth. This is what one would expect in a world dominated by fat-tail firm level demand (or quality) shocks. An industry grows because some of its firm draw a large positive shock. This mechanically leads to higher concentration. A doubling of the CR4 is correlated with a 13% increase in labor productivity. Wages rise by only 3% implying that productivity gains are not passed on to workers. Unit labor costs, therefore, fall by 10%. In a competitive economy, this would lead to lower prices and increased welfare for consumers. However, prices remain flat – implying a 11% increase in mark-ups[29].

[29] Our results are fairly consistent withGanapati (2018). Using Table 4 of Ganapati (2018), we obtain a regression beta between mark-up increases and concentration of 0.05 for non-manufacturing, compared to 0.08 in our data:

Table 8: *Concentration and Mark-up Decomposition: Granular Industries*

Table shows the results of industry-level OLS regressions of contemporaneous 5-year changes in concentration, mark-ups, prices and ULC for as long as data are available. Observations are unweighted to mirror Ganapati (2018). Standard errors in brackets, clustered at industry-level. + p<0.10, * p<0.05, ** p<.01.

Panel A. All Industries

	(1) Mark-ups	(2) Prices	(3) ULC	(4) Wages	(5) LP
s5logcr4	0.12**	0.03*	-0.09**	0.03**	0.13**
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Cons	0.04**	0.10**	0.06**	0.16**	0.10**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sec x Yr FE	Y	Y	Y	Y	Y
R2	.25	.21	.15	.18	.12
Observations	2,083	2,084	2,083	2,083	2,083

Panel B. Manufacturing

	(1) Mark-ups	(2) Prices	(3) ULC	(4) Wages	(5) LP
s5logcr4	0.13**	0.04**	-0.10**	0.04**	0.14**
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Cons	0.05**	0.10**	0.04**	0.15**	0.11**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sec x Yr FE	Y	Y	Y	Y	Y
R2	.24	.2	.11	.11	.091
Observations	1,682	1,682	1,682	1,682	1,682

Panel C. Non-Manufacturing

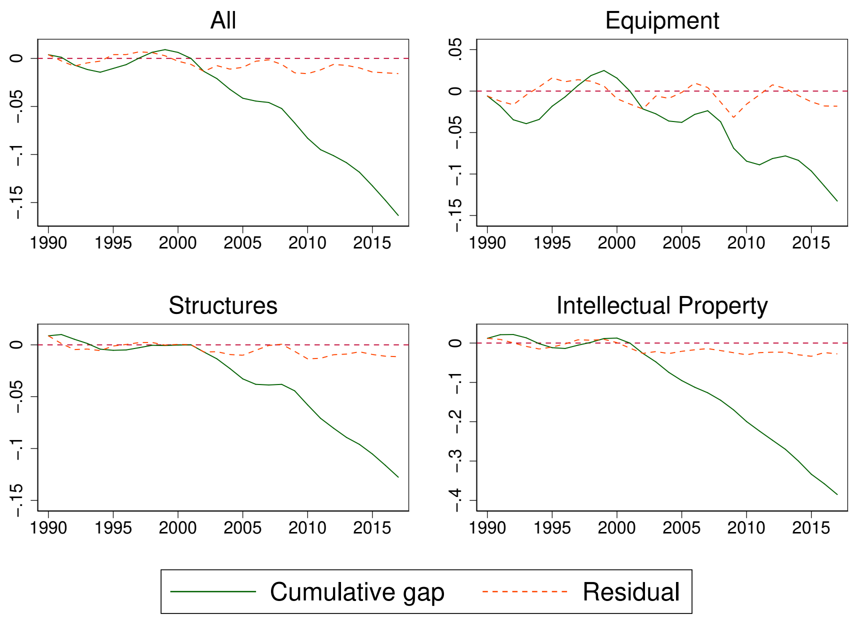
	(1) Mark-ups	(2) Prices	(3) ULC	(4) Wages	(5) LP
s5logcr4	0.08**	-0.01	-0.09**	0.00	0.10**
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Cons	0.00	0.11**	0.11**	0.17**	0.06**
	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
Sec x Yr FE	Y	Y	Y	Y	Y
R2	.15	.26	.18	.35	.17
Observations	401	402	401	401	401

$$\beta_{\mu,CR4} = \beta_{p,CR4} - \beta_{w,CR4} + (\beta_{q,CR4} - \beta_{N,CR4})$$
$$0.05013 = -0.00421 - \underbrace{[0.00596 - (0.0477 - (-0.0126))]}_{-0.05434}$$

B.3 Investment.

In figure 24 we show the residual and cumulative gap from the regression $K_t = \beta_0 + \beta_1 Q_{t-1} + \epsilon_t$, where Q represents Tobin's Q. We run this regression for the entire capital stock and also for the three types of capital reported in BEA's fixed asset tables: Equipment, Structures and Intellectual Property.

Figure 24: *Growth Rates of Capital Stock vs Predicted by Q-theory*



Notes: Annual Data. Growth rate of private nonresidential fixed assets; based on section 4.2 of the BEA's fixed assets tables. Q for Non Financial Business sector from US Flow of Funds accounts.

In order to confirm that changes in the profit rate of leaders is not only a between-firms effect but also within-firms, we estimate

$$\left(\frac{OIADP}{SALE}\right)_{i,j,t} = \beta_t \times Lead_{i,j,t} + \delta_i + \gamma_t + \varepsilon_{jt}, \tag{7}$$

where $Lead_{ijt}$ is an indicator equal to one for firms in the top quantile of the market value distribution, by industry; while δ_i and γ_t denote firm and year fixed effects, respectively. Observations are weighted by sales. Coefficient γ_t captures the average within-firm change in profits, while β_t captures an incremental effect for leaders firms. We plot $\beta_t + \gamma_t$ as the total effect on leaders.

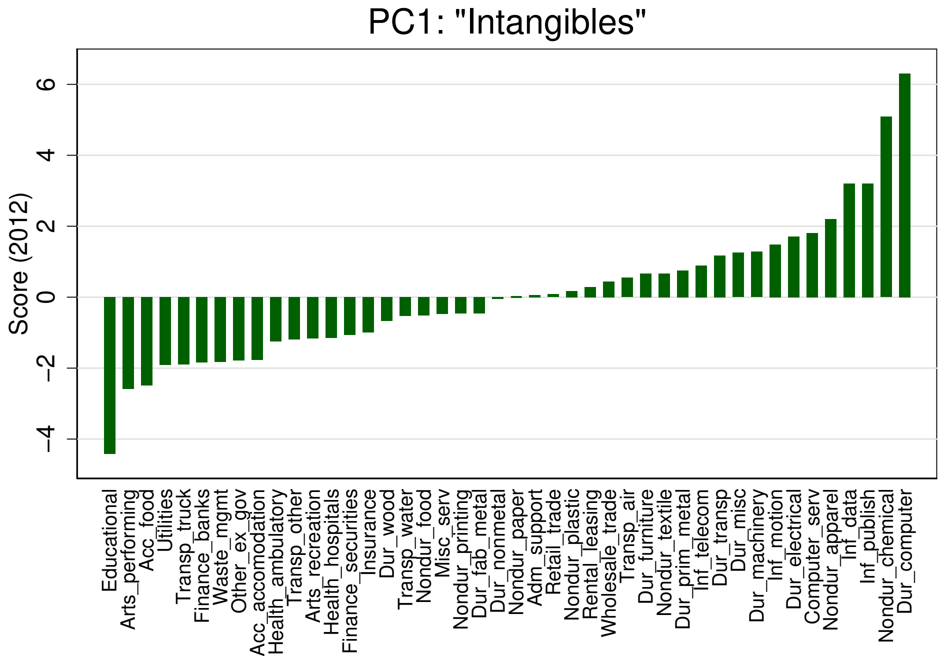
Figure 25: *Within-firm Change in Profit Margin for Leaders vs. Laggards*

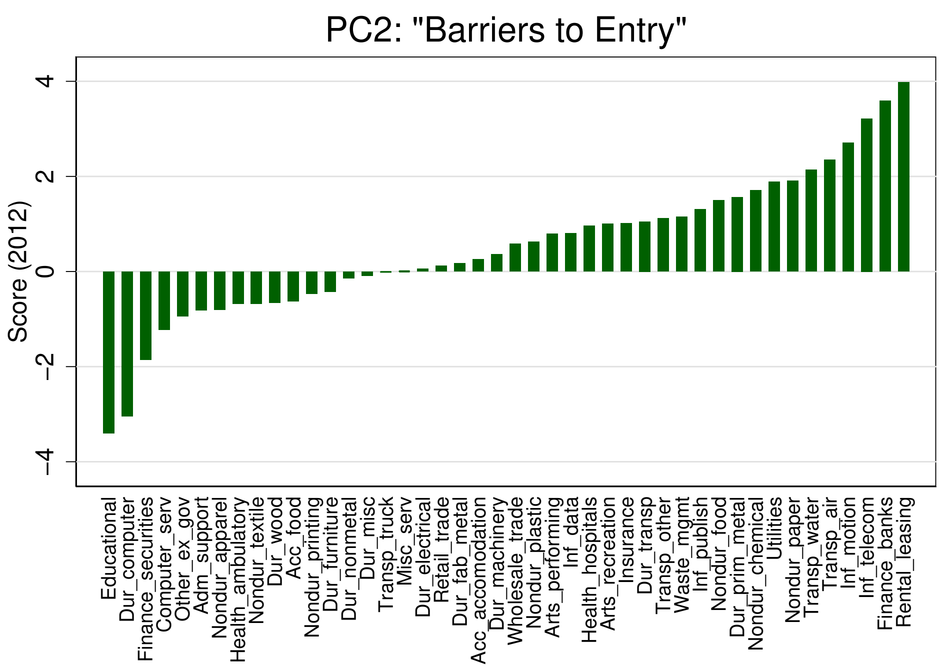
Notes: Compustat NA. Figure plots the estimated within-firm change in profits for leaders and laggards, following equation 7. See text for details.

C Appendix for Section 4: PCA

Figure 26 shows the loadings on PC1 and PC2, as of 2012, for each industry.

Figure 26: *Principal Component Scores, by Industry*





Notes: see text for details and data appendix for variable definitions.

D Model Appendix

D.1 Demand System

There is a continuum of industries indexed by $j \in [0, 1]$ and a continuum of firms $i \in [0, N_{j,t}]$ in each industry. A particular firm is therefore indexed by (i, j), i.e., i'th firm in industry j.

Firms' outputs are aggregated at the industry level as

$$Y_{j,t}^{\frac{\sigma_j - 1}{\sigma_j}} = \int_0^{N_{j,t}} h_{i,j,t}^{\frac{1}{\sigma}} \left(y_{i,j,t}\right)^{\frac{\sigma_j - 1}{\sigma_j}} di$$

where σ is the elasticity between different firms in the same industry and h are firm-level demand shocks, with a mean of 1. Industry outputs are aggregated into a final consumption bundle

$$\bar{Y}_t = \int_0^1 H_{j,t}^{\frac{1}{\epsilon}} Y_{j,t}^{\frac{\epsilon - 1}{\epsilon}} dj$$

where ϵ is the elasticity of substitution between industries. This demand structure implies that there exists an industry price index

$$P_{j,t}^{1-\sigma_j} \equiv \int_0^{N_{j,t}} h_{i,j,t} p_{i,j,t}^{1-\sigma_j} di$$

such that the demand for good i is given by

$$y_{i,j,t} = h_{i,j,t} Y_{j,t} \left(\frac{p_{i,j,t}}{P_{j,t}}\right)^{-\sigma_j}$$

Similarly, there exists an aggregate price index

$$\bar{P}_t^{1-\epsilon} \equiv \int_0^1 H_{j,t} P_{j,t}^{1-\epsilon} dj$$

such that industry demand is

$$Y_{j,t} = H_{j,t} \bar{Y}_t \left(\frac{P_{j,t}}{\bar{P}_t}\right)^{-\epsilon}$$

D.2 Production

The production function of firm i, j is Cobb-Douglass

$$y_{i,j,t} = a_{i,j,t} k_{i,j,t}^{\alpha_j} n_{i,j,t}^{1-\alpha_j}$$

and there is a fixed cost of production ϕ_j. Firms take the wage W and the rental rate R as given when they hire capital and labor. The Cobb-Douglass function, like any CRS function, leads to a constant marginal cost

$$\chi_{i,j,t} = \frac{1}{a_{i,j,t}} \left(\frac{R_t}{\alpha_j}\right)^{\alpha_j} \left(\frac{W_t}{1-\alpha_j}\right)^{1-\alpha_j}$$

Cost minimization implies that all firms choose the same (optimal) capital labor ratio

$$\frac{\alpha_j}{1-\alpha_j}\frac{n_{i,j,t}}{k_{i,j,t}}=\frac{R_t}{W_t}$$

The average cost is $\chi_{i,j,t}y_{i,j,t}+\phi_j$

Profits are

$$\pi_{i,j,t}=p_{i,j,t}y_{i,j,t}-\chi_{i,j,t}y_{i,j,t}-\phi_j$$

If we define the mark-up of price over marginal cost

$$p_{i,j,t}\equiv\left(1+\mu_{i,j}\right)\chi_{i,j,t}$$

Then profits are

$$\begin{aligned}
\pi_{i,j,t} &= \frac{\mu_{i,j}}{1+\mu_{i,j}}p_{i,j,t}y_{i,j,t}-\phi_j \\
&= h_{i,j,t}\frac{\mu_{i,j}}{\left(1+\mu_{i,j}\right)^{\sigma}}\left(\chi_{i,j,t}\right)^{1-\sigma_j}P_{j,t}^{\sigma_j}Y_{j,t}-\phi_j \\
&= h_{i,j,t}\frac{\mu_{i,j}}{1+\mu_{i,j}}\left(\frac{1+\mu_j}{1+\mu_{i,j}}\frac{a_{i,j,t}}{A_{j,t}}\right)^{\sigma_j-1}P_{j,t}Y_{j,t}-\phi_j
\end{aligned}$$

where $A_{j,t}$ is industry-average productivity and μ_j is industry-average mark-up.

Nominal revenues are

$$p_{i,j,t}y_{i,j,t}=p_{i,j,t}^{1-\sigma_j}h_{i,j,t}P_{j,t}^{\sigma_j}Y_{j,t}$$

and the market share is

$$s_{i,j,t}=\frac{p_{i,j,t}y_{i,j,t}}{P_{j,t}Y_{j,t}}=\frac{h_{i,j,t}}{N_j}\left(\frac{\left(1+\mu_j\right)a_{i,j,t}}{\left(1+\mu_{i,j}\right)A_{j,t}}\right)^{\sigma_j-1}$$

E Data Appendix

We use a wide range of aggregate-, industry- and firm-level data, summarized in Table 9 and described in the rest of this section. We begin by describing the three datasets used repeatedly throughout the paper: Compustat North America, Compustat Global and US Economic Census Concentration Ratios (section E.1). We then discuss how these, and the remaining datasets are used to generate specific results.

E.1 Main dataset

E.1.1 Compustat North America

Sample Selection. Our primary firm-level data is based on tables Funda, Company and Exrt_mth from Compustat North America, obtained via WRDS. Compustat North America includes all public and some private firms in North America. Data are available from 1950, but coverage is fairly thin until the 1970s. We apply standard screens (consol = "C", indfmt = "INDL", datafmt = "STD", popsrc = "D"), and ex-

Table 9: *Summary of Key Data Sources*

	Source	Key Data fields	Granularity
Sectoral	Financial Accounts of the United States via FRED	I, K, OS, ...	Sector (NFCB, NFNCB)
Industry	OECD STAN	OS, $PROD$	ISIC L2
	EU KLEMS 2018	LS	~ISIC L2
	BEA GDP by Industry	Output & prices	~NAICS L3 (summary) and ~NAICS L4 (detailed)
	BEA Fixed Assets Tables	I, K	~NAICS L3
	BLS Multifactor Productivity Tables	TFP, P, Q, ...	~NAICS L3
	Economic Census	Concentration	NAICS L3-L6
	NBER-CES database	P, Q	NAICS L6
	Peter Schott's website	Imports, NTR Gap	NAICS L6
Firm	Compustat (NA and Global)	Q, I, K and OS	Firm
	Peters & Taylor	Intangible K	Firm

clude firm-year observations with missing year, sales, assets, or $gvkey$.[30] We use the exchange rates in exrt_mth to convert all financials to USD. We keep all firms for our global analyses, but restrict the sample to US-headquartered firms with USD currency codes for US-specific analyses (LOC = "USA", CURCD = "USD"). We complement Compustat with the firm-level intangible capital estimates of Peters and Taylor (2016) (WRDS table total_q); and use CRSP table msf as well as the CRSP-Compustat linking table (ccmxpf_linktable) to fill in missing stock prices in Compustat, when needed (see replication file for details).

Industry Segments. We use the industry codes in the Compustat Company table. NAICS codes are populated for all firms that existed after 1985, but are sometimes missing for firms that exited beforehand. We map those firms to the most common NAICS-4 industry among those firms with the same SIC code and non-missing NAICS. We also map all retired/new NAICS codes from the 1997, 2002 and 2012 versions to NAICS 2007 using the concordances in link.

We then map NAICS codes to BEA and EU KLEMS industries. For BEA industries, we use the mapping in tab 'NAICS codes' of file GDPbyInd_GO_1947-2017.xls. This includes 63 granular industries. We group 'Motor vehicles, bodies and trailers, and parts' and 'Other transportation equipment', and keep only 'Hospitals and Nursing' (which groups 'Hospitals' and 'Nursing and Residential Care facilities') because only the grouped industries are covered in the BLS' multifactor tables. We exclude Real Estate given the 2000's boom, as well as 'Management of companies and enterprises' because there are no companies in Compustat that map to this category. This leaves 59 industry groupings, summarized in Table 10. Firms with NAICS codes 999 cannot mapped to BEA industries. These firms are mapped to an 'other' industry, which is included in those analyses that do not rely on aggregate data.

EU KLEMS (and STAN) industries follow the ISIC Rev. 4 hierarchy. We map firms from NAICS 2007

[30] We also address selected data issues manually (e.g., outliers in sales growth, especially when reported currency changes). See replication code for details.

to ISIC Rev. 4 using the concordance available at link as follows: first, we map each NAICS-6 segment to the most common ISIC Level 2 segment (by number of mappings) based on the the concordance. This mapping is one-to-one for most NAICS-6 segments; and for the remaining segments there is usually a single most common ISIC Level 2 segment. For the few cases where NAICS-6 segments map with equal likelihood to more than one ISIC Level 2 segment, we follow the same methodology but with NAICS-5 codes (and so on).[31] We then map each ISIC Rev. 4 Level 2 segments to the 27 EU KLEMS industries.

Concentration Ratios. We use the resulting dataset to compute Compustat-based concentration ratios. Compustat coverage as a share of the economy varies over time (as more firms go public) and across industries (depending on the nature of production); and the importance of foreign competition varies over time. To ensure CRs are stable over time and across industries, and account for imports we compute:

$$CR4_{jt} = \frac{\sum\limits_{i \in \{j, \text{top4}\}} sale_{it}^{CPSTAT}}{s_{jt}^{CPSTAT}} \times c_{jt}^{MA}$$

where s_{it}^{CPSTAT} denotes sales for firm i which belongs to industry j and s_{jt}^{CPSTAT} denotes sales across all Compustat firms in industry j. c_{jt}^{MA} denotes the coverage adjustment, equal to a three-year centered moving average of the yearly coverage ratio ($c_{jt} = \frac{s_{jt}^{CPSTAT}}{s_{jt}^{BEA} + Imports_{jt}}$, where s_{jt}^{BEA} denotes gross output from the BEA and Imports denotes imports from Peter Schott's data). We use a moving average to smooth the impact of FX volatility given that Compustat sales include both domestic and foreign sales. c_{jt} can exceed 1 for exporting industries and may be affected by FX volatility even if 'real output' coverage remains flat, so we cap c_{jt}^{MA} at 1.25 (which assumes slightly higher domestic CR relative to global CRs). Last, to ensure the estimated CRs are robustly estimated, we include only industries where average database coverage after 2000 exceeds 10%. See replication code for details.

Other Definitions.

- **Market Value of Equity (ME):** ME is defined as the total number of common shares outstanding (item CSHO) times the closing stock price at the end of the fiscal year (item PRCC_F). When either CSHO or PRCC_F are missing in Compustat, we fill-in the value using CRSP. If ME is also missing in CRSP, we use PRCC_C x CSHO.

- **Market Value (MV)**: MV is defined as the market value of equity (ME) plus total liabilities (LT) and preferred stock (PSTK)

- Q**:** firm-level Q is defined as the ratio of market value to total assets (AT). We cap Q at 10 and winsorize it at the 2% level, by year to mitigate the impact of outliers. See Gutiérrez and Philippon (2017b) for a discussion of alternate definitions of Tobin's Q.

[31] In some cases, Compustat NAICS codes contain fewer than six digits. In that case, we repeat the process using NAICS-5 to NAICS-2 codes. Firms that cannot be mapped to an ISIC segment (those with NAICS code 999 are excluded from industry-level analyses).

Table 10: *Mapping of BEA industries to segments*

BEA code	BEA Industry	Mapped segment
	Agriculture, forestry, fishing, and hunting	Omitted
1100	Farms	Agr_farm
1130	Forestry, fishing, and related activities	Agr_forest
	Mining	Omitted
2110	Oil and gas extraction	Min_oil_and_gas
2120	Mining, except oil and gas	Min_ex_oil
2130	Support activities for mining	Min_support
2200	**Utilities**	Utilities
2300	**Construction**	Construction
	Durable goods manufacturing	Omitted
3210	Wood products	Dur_wood
3270	Nonmetallic mineral products	Dur_nonmetal
3310	Primary metals	Dur_prim_metal
3320	Fabricated metal products	Dur_fab_metal
3330	Machinery	Dur_machinery
3340	Computer and electronic products	Dur_computer
3350	Electrical equipment, appliances, and components	Dur_electrical
3360	Motor vehicles, bodies and trailers, and parts	Dur_transp
3360	Other transportation equipment	Dur_transp
3370	Furniture and related products	Dur_furniture
3390	Miscellaneous manufacturing	Dur_misc
	Nondurable goods manufacturing	Omitted
3110	Food and beverage and tobacco products	Nondur_food
3130	Textile mills and textile product mills	Nondur_textile
3150	Apparel and leather and allied products	Nondur_apparel
3220	Paper products	Nondur_paper
3230	Printing and related support activities	Nondur_printing
3240	Petroleum and coal products	Nondur_petro
3250	Chemical products	Nondur_chemical
3260	Plastics and rubber products	Nondur_plastic
4200	**Wholesale trade**	Wholesale_trade
4400	**Retail trade**	Retail_trade
	Transportation and warehousing	Omitted
4810	Air transportation	Transp_air
4820	Railroad transportation	Transp_rail
4830	Water transportation	Transp_water
4840	Truck transportation	Transp_truck
4850	Transit and ground passenger transportation	Transp_passenger
4860	Pipeline transportation	Transp_pipeline
4870	Other transportation and support activities	Transp_other
4930	Warehousing and storage	Transp_storage

Table 10: Mapping of BEA industries to segments (cont'd)

BEA code	Sector/Industry	Mapped industry
	Information	Omitted
5110	Publishing industries (includes software)	Inf_publish
5120	Motion picture and sound recording industries	Inf_motion
5130	Broadcasting and telecommunications	Inf_telecom
5140	Information and data processing services	Inf_data
	Finance and insurance	Omitted
5210	Federal Reserve banks	Finance_banks
5210	Credit intermediation and related activities	Finance_banks
5230	Securities, commodity contracts, and investments	Finance_securities
5240	Insurance carriers and related activities	Insurance
5250	Funds, trusts, and other financial vehicles	Finance_funds
	Real estate and rental and leasing	Omitted
5310	Real estate	Omitted
5320	Rental and leasing services and lessors of intangible assets	Rental_leasing
	Professional, scientific, and technical services	Omitted
5411	Legal services	Legal_serv
5415	Computer systems design and related services	Computer_serv
5412	Miscellaneous professional, scientific, and technical services	Misc_serv
5500	**Management of companies and enterprises**	Omitted
	Administrative and waste management services	Omitted
5610	Administrative and support services	Adm_support
5620	Waste management and remediation services	Waste_mgmt
6100	**Educational services**	Educational
	Health care and social assistance	Omitted
6210	Ambulatory health care services	Health_ambulatory
6220	Hosp and nursing	Health_hospitals
6220	Hospitals	Omitted
6220	Nursing and residential care facilities	Omitted
6240	Social assistance	Health_social
	Arts, entertainment, and recreation	Omitted
7110	Performing arts, spectator sports, museums, and related activities	Arts_performing
7130	Amusements, gambling, and recreation industries	Arts_recreation
	Accommodation and food services	Omitted
7210	Accommodation	Acc_accomodation
7220	Food services and drinking places	Acc_food
8100	**Other services, except government**	Other_ex_gov

- **Total Capital** (K^{PT})**:** K^{PT} is set equal to PPEGT plus K_INT, where the former is included in Compustat and the latter is provided by Peters and Taylor (2016).

- **Firm Age**: Firm age is defined as the number of years over which a firm appears in Compustat, irrespective of whether the underlying data fields satisfy our exclusion restrictions (i.e., we measure age before imposing any exclusion restrictions).

- **Ratios**: We also compute a variety of ratios as described in the text (e.g., SALE/COGS, XSGA/XOPR). All of these ratios are winsorized at the 2% and 98% level, by year to mitigate the impact of outliers.

E.1.2 Compustat Global

Global concentration measures are based on Compustat Global, which includes most public firms across advanced economies. Data are available from 1987, but coverage is fairly thin until the late-1990s. We download tables g_funda, g_company and g_exrt_mth via WRDS. We apply the same screens as for the US (consol = "C", indfmt = "INDL", datafmt = "STD", popsrc = "I") and exclude firm-year observations with missing year, sales, assets, or $gvkey$.[32] We use the exchange rates in exrt_mth to convert all financials to USD. For a few firms, currency codes and financials appear inconsistent – particularly when currency codes change. We therefore drop firms (gvkeys) entirely whenever sales or assets increase or decrease by a factor of 20 in the same year as the currency code changes. Firms are mapped to countries/regions using headquarter location (LOC). We then use the same definitions and mapping procedure as for the US.

E.1.3 Economic Census Concentration Ratios

Last, we obtain sales, employment and payroll data by industry from the US Economic Census' Concentration accounts. The data include breakdowns for the top 4, 8, 20 and 50 firms in each industry along with industry totals, and are published every five years. All firms operating within a given SIC/NAICS category in the United States are included. See link for additional details.

Data before 1992 is based on the SIC system. For manufacturing, we use the retrospective tabulation based on unified SIC codes published in the 1992 Economic Census. For non-manufacturing, we use the data as reported, which follows the 1987 SIC system in both 1987 and 1992, though there are small adjustments across years. Data after 1997 is based on NAICS, with each of the 1997, 2002, 2007 and 2012 reports using slightly different NAICS vintages. Like Ganapati (2018), we restrict our sample to consistently defined SIC/NAICS codes over each five-year period. Data for service industries are reported by tax-paying segments. We keep tax-payable firms because they are reported consistently over time and are closest to our analysis. Data for wholesale trade are reported as a total and by type of merchant (e.g., merchant wholesaler, manufacturer). We keep only the total.

Table 3 shows the coverage of the data. We restrict our sample to the post-1987 period, when concentration increased. There is continuous coverage for the manufacturing sector over the entire time period at the 4-digit SIC and 6-digit NAICS levels. Coverage for non-manufacturing sectors is spottier. Wholesale trade,

[32] We also address selected data issues manually (e.g., outliers in sales growth, especially when reported currency changes). See replication code for details.

retail trade and services are covered since 1987, as well as some transportation and communication sectors. All major industries except agriculture, mining and construction are covered after 1997.

We use these data in four ways: first, we use the reported concentration ratios directly in some of our figures and/or regressions. Second, we compute census-based import-adjusted concentration as

$$CR8_{jt}^{IA} = CR8_{jt} \times \frac{sale_{jt}}{sale_{jt} + imp_{jt}} = CR8_{jt} \times \text{US Share}_{jt}$$

where $CR8_{jt}$ and $sale_{jt}$ are based on the US Economic Census; and $imports_{jt}$ is based on Peter Schott's data (set to zero when missing). Third, we aggregate census concentration ratios to BEA industries since 1997, for use in the PCA analysis. Census concentration measures follow the NAICS hierarchy, which almost always maps one-to-one to BEA industries. When this is not the case, we first aggregate (domestic) concentration ratios to BEA industries by taking a sales-weighted average; and then apply the import adjustment. For some regressions, we interpolate Census concentration measures between economic census years. Last, we combine the concentration data with price indices from the NBER-CES database for manufacturing and the BEA's detailed GDP by Industry accounts for our analyses of productivity and prices. See below for details.

E.2 Details on the Construction of Results

E.2.1 Introduction

Figures 2, Panel A: Profits. Profits rates are based on OECD table STANI4_2016, which follows ISIC Rev. 4 segments. Data are available for 37 countries. We focus on the nonagriculture business sector excluding real estate (D05T82X), and include only advanced economies for which gross profits data are available since 2000: the EU28 ex. BGR, CYP, HRV, MLT, ROU plus JPN, KOR, NOR, and the USA. AUS, CHE and CAN are excluded because data are available after 2005. We convert all nominal quantities to US dollars using the OECD's exchange rates, available at link. We define the gross profit rate as the ratio of GOPS to PROD. We aggregate across countries by taking the production-weighted average.

Figure 2, Panel B. Concentration. We then measure concentration using the same calculation as for the US, with three exceptions: first, we do not adjust for imports. Second, we use the 27 industries defined in EU KLEMS, instead of BEA industries. Third, we use gross output data from OECD STAN to adjust for Compustat coverage, instead of BEA gross output data. To ensure consistency between STAN output and Compustat sales, we drop firms in country x industry x years where STAN data are not available. This means our EU-wide series includes 23 countries (EU28 ex Bulgaria, Croatia, Cyprus, Malta, Romania). Concentration is measured at the region x industry-level. We then compute changes since 2000, and aggregate across industries within a region, weighing by production in constant 2009 prices (STAN item PRDK). We use constant prices because variations in oil prices can introduce undue volatility to the weights of petroleum-dependent industries (see Jones et al. (2019)).

Figure 2, Panel C. Labor Share. Figures reports the value-added weighted average change in the labor share for the Market Economy based on EU KLEMS (KLEMS LAB/VA). Data for most countries are available since 1995, but we include countries for which data are available at least since 2000. Thus, the EU series includes EU28 ex. HRV, HUN, MLT, POL. We then compute changes since 2000, and aggregate across industries within a region, weighing by value added (EU KLEMS item va).

E.2.2 Measurement Issues

Figure 6: Mark-up vs Profits. GOS/PROD for nonagriculture business sector excluding Real Estate from OECD STAN, as described above. Compustat series equal to the sales-weighted average of SALE/COGS across all Compustat firms in a given year x region, included in sample above. Data reported for EU since 1989, but note that a sizable portion of European firms report COGS only after ∼2005.

Table 7: Summary of Income Statement. Start from US Compustat sample described above. Keep firm x year pairs for neither SALE, COGS, SG&A, OIBDP, DP and OIADP are missing. Report the sales weighted average of the ratio of COGS/SALE, SG&A/SALE, etc across all firms and years in a given decade. All ratios are winsorized at the 2% level by year.

Figure 16: SALE/COGS vs. SG&A intensity for high-mark-up firms. Start from US Compustat sample described above. Drop firms with missing SALE/COGS or XSGA/XOPR. Identify firms in the top 25th percentile of the SALE/COGS distribution. Report a scatter plot of the sales-weighted average ratio of SALE/COGS and XSGA/XOPR across those firms, in each year. As above, SALE/COGS and XSGA/XOPR are winsorized at the 2% level by year.

Figure 4 and 5. All the analyses of mark-up measurement using the China Shock are based on NAICS-6 manufacturing industries. We complement Compustat with three additional datasets:

- **Import and Exports**: Import and export data are sourced from Peter Schott's website and was first used in Schott (2008). Data are available by HS-code x year from 1989 to 2017, but include a mapping to NAICS-6 industries which follows the concordance of Pierce and Schott (2012). We use these data to estimate import penetration and import-adjusted concentration at different levels of granularity (NAICS-6 as well as BEA industries).

- **NTR gap:** We also gather Non-Normal-Trade-Relations tariff gaps from the replication file of Pierce and Schott (2016). NTR gaps are defined for NAICS level 6 industries.[33]

- **NBER-CES database:** Last, we use the NBER-CES database, which includes output and productivity data by NAICS Level 6 manufacturing industry from 1971 to 2011. It also includes measures of the production structure in each industry (such as production workers as a share of total employment, the log average wage, etc.), which are used as controls in regressions and to test alternate theories of concentration.

[33] NTR gaps are available in file 'gaps_by_naics6_20150722_fam50', which includes NTR gaps for each NAICS Level 6 code.

These datasets are merged into the main Compustat sample by NAICS-6 industry x year, which includes the total capital estimates of Peters and Taylor (2016). See main text for details on the construction of each result.

E.2.3 Entry, Exit and Turnover

All figures are based on our main Compsutat sample described above. See text for details.

E.2.4 Joint evolution of Concentration, TFP and prices

Table 2: Concentration, TFP, Prices and Mark-ups: BLS industries. Merge Compustat import-adjusted concentration measures with BLS KLEMS data on prices and productivity. Compute mark-ups and implement regression.

Table 3: Concentration vs Prices: Detailed industries. For manufacturing, we merge Economic Census concentration ratios with sales, prices, employment and payroll data from the NBER-CES database. The data are based on 4-digit SIC codes before 1997 and 6-digit NAICS after 1997. For non-manufacturing, merge sales, payroll, employment and concentration data from the Economic Census to prices from the BEA's detailed GDP by Industry accounts (files GDPbyInd_GO_NAICS_1997-2016.xlsx and GDPbyInd_GO_SIC.xlsx). These files include $\sim$400 industries, with more than 200 corresponding to manufacturing industries. Ganapati (2018) uses more detailed accounts, but we focus on this higher level of aggregation because, even for these accounts, the BEA acknowledges that "the more detailed estimates are more likely to be either based on judgmental trends, on trends in the higher-level aggregate, or on less reliable source data." Some of the BEA industries aggregate several NAICS codes. We manually map as many codes as possible, and aggregate concentration ratios by taking a weighted average when needed. We then compute quantities, labor productivity and mark-ups as defined in the text – and estimate the regressions.

E.2.5 Investment and Profits by Leaders vs. Laggards

Table 4: Investment, Capital and Profits by Leaders and Laggards. Rank firms by market value. Define a firm as leader if it is the top firm in a given industry or the cumulative market value up to and including this firm is below 33% of the industry market value. Repeat the exercise for mid-performers (33-66% of MV) and the bottom 33%. Next, compute the total OIBDP, CAPX + R&D, PP&E and Capital K (including intangibles as estimated by Peters and Taylor (2016)) by year and by MV group x year. Estimate the share of a given measure – say OIBDP – as the ratio of leader OIBDP to total OIBDP in a given year. Because firms are discrete, the actual share of market value in each grouping varies from year to year. To improve comparability, re-scale shares by the ratio of 33.33% to the share of market value. Report the average across all years in a given period.

Table 5: Investment by Leaders We start from our base Compustat sample, mapped to BEA industries. Deflate capital stock using the industry-level price of capital reported in the BEA's fixed assets tables (see

below for a description). Compute yearly change in (deflated) capital stock and winsorize at the 2% and 98% level by year. Include only firm-year pairs with non-missing PPEGT, K_INT and K_PT. Define leaders as firms with market value in the top quantile by BEA industry and year. Estimate regression as reported in the text.

E.2.6 PCA

Our PCA analysis is based on the BEA industries described in Table 10. We define the data sources and definitions for all measures included in the analysis. The rest of the details are provided in the text.

- **Census Concentration (cr4_cen and Dcr4_cen):** The level in census concentration, as described in Section E.1 as well as the change since 2007
- **BEA Intangible Capital Share (intan_kshare_bea and Dintan_kshare_bea):** ratio of intellectual property capital to total capital as measured in Section 3 of the BEA Fixed Assets tables, available at link; as well as the change since 1997
- **Intangible Capital Share (intan_kshare_med_pt):** Define the firm-level intangible capital share as the ratio of internally-developed intangibles K_INT - INTAN (from Peters and Taylor (2016)) to total capital (K_INT + PPEGT). Compute the median across all firms in a given industry x year. Similar results including externally developed intangibles.
- **Import share (import_share)**: ratio of imports from Peter Schott's data to the sum of gross output and imports.
- **BEA Profit Margin (profit_margin_bea):** ratio of net operating surplus to gross output as measured by the BEA's GDP by Industry accounts (file GDPbyInd_GO_1947-2017).
- **Compustat Median Profit Margin (profit_margin_med_cp)**: Define firm-level profit margin as the ratio of operating income after depreciation to sales (OIADP/SALE). Compute the median across all firms in a given industry x year.
- **US KLEMS inputs:**
  - Labor Share (ls_kl) defined as the ratio of total labor expenses to gross output minus intermediate inputs.
  - TFP growth (dtfp_kl) equals the five-year log-change in a given industry's multifactor productivity index (MFP)
  - Price, ULC and Mark-up growth (Dlogp_kl, Dlogulc_kl and Dlogmu_kl, respectively) defined as described in section E.2.4 above.
- **Leader Turnover (lead_turnover_mv):** market-value based turnover rate, as defined in section XX above.

- **Compustat firm-level leader investment gap (ikgap_cp):** we roughly follow Crouzet and Eberly (2018). Define the net investment rate for firm i in industry j as the log-change in (deflated) total capital, $\Delta \log(K_{ijt}^{PT})$, using the industry-level deflator from the BEA's fixed assets tables. Then, estimate $\Delta \log(K_{ijt}^{PT}) = \beta \bar{Q}_{jt} + \beta_2 \log(Age_{ijt}) + \delta_i + \gamma_t + \varepsilon_{ijt}$, where we control for firm-age, industry average Q as well as firm and year fixed effects. The year fixed effects measure the annual investment gap.

Table 11: *Mapping of BEA industries to segments*

	1987	1992	1997	2002	2007	2012
Sector	SIC		NAICS			
Agriculture						
Mining						
Construction						
Manufacturing	X	X	X	X	X	X
Transportation	Partial		X	X	X	X
Communication	Partial		X	X	X	X
Utilities				X	X	X
Wholesale Trade	X	X	X	X	X	X
Retail Trade	X	X	X	X	X	X
FIRE				X	X	X
Services	X	X	X	X	X	X