Declining Competition and Investment in the U.S.*

Germán Gutiérrez† and Thomas Philippon‡

November 2017

Abstract

Since the early 2000's, US industries have become more concentrated and profitable while non residential business investment has been weak relative to fundamentals. The interpretation of these trends is controversial. We test four explanations: decreasing domestic competition (DDC); increases in the efficient scale of operation (EFS); intangible investment (INTAN); and globalization (GLOBAL). We first present new evidence that supports DDC against EFS: concentration rose in the U.S. but not in Europe; the relationship between concentration and productivity was positive in the 1990s, but is zero or negative in the 2000s; and industry leaders cut investment when concentration increased. We then establish the causal impact of competition on investment using three identification strategies: Chinese competition in manufacturing; noisy entry in the late 1990s; and discrete jumps in concentration following large M&As. Taking into account INTAN and GLOBAL, we find that more (less) competition causes more (less) investment, particularly in intangible assets by industry leaders. We conclude that DDC has resulted in a shortfall of non residential business capital of 5 to 10% by 2016.

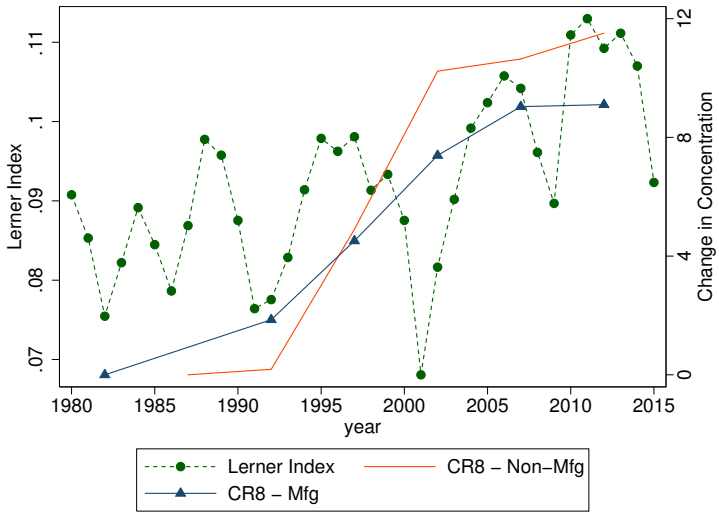
*We are grateful to Bob Hall, Janice Eberly, Steve Davis and Christopher House for their comments and discussions, as well as Viral Acharya, Manuel Adelino, Olivier Blanchard, Ricardo Caballero, Charles Calomiris, Alexandre Corhay, Emmanuel Farhi, Jason Furman, Xavier Gabaix, John Haltiwanger, Campbell Harvey, Glenn Hubbard, Ron Jarmin, Boyan Jovanovic, Sebnem Kalemli-Ozcan, Ralph Koijen, Howard Kung, Javier Miranda, Holger Mueller, Valerie Ramey, David Robinson, Tano Santos, René Stulz, Alexi Savov, Philipp Schnabl, Jose Scheinkman, Martin Schmalz, Lukas Schmid, Carolina Villegas-Sanchez, Johannes Wieland, Luigi Zingales, and seminar participants at NYU, ESSIM, Columbia University, University of Chicago, Harvard, Duke, NBER EFG and NBER ME meetings for stimulating discussions.

†New York University. ggutierr@stern.nyu.edu

‡New York University, CEPR and NBER. tphilipp@stern.nyu.edu

Two important stylized facts have emerged in recent years regarding the U.S. business sector. The first fact is that concentration and profitability have increased across most U.S. industries, as shown by Grullon et al. (2016). Figure 1 shows the aggregate Lerner index (operating income over sales) across all Compustat firms along with the change in weighted average 8-firm concentration ratio in manufacturing and non-manufacturing industries.[1]

Figure 1: Concentration and Mark-ups



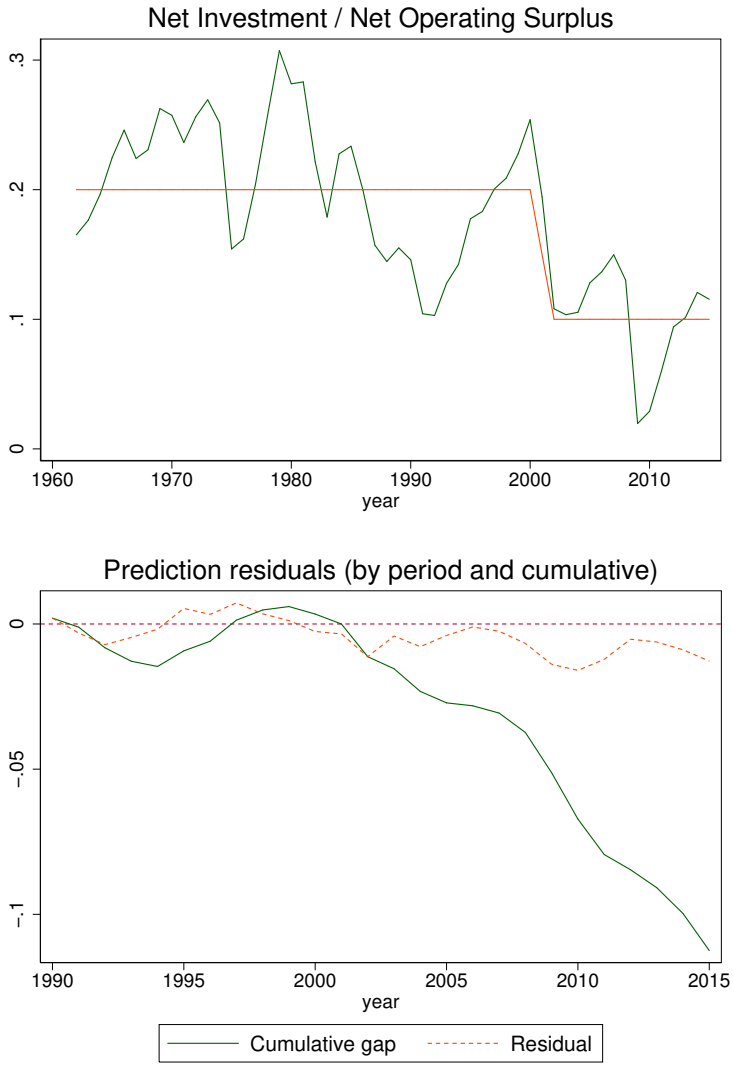
Notes: Lerner Index from Compustat, defined as operating income before depreciation minus depreciation divided by sales. 8-firm CR from Economic Census, defined as the market share (by sales) of the 8 largest firms in each industry. Data before 1992 based on SIC codes. Data after 1997 based on NAICS codes. Data for Manufacturing reported at NAICS Level 6 (SIC 4) because it is only available at that granularity in 1992. Data for Non-Manufacturing based on NAICS level 3 segments (SIC 2).

The second stylized fact is that business investment has been weak relative to measures of profitability, funding costs, and market values since the early 2000s. The top chart in Figure 2 shows the ratio of aggregate net investment to net operating surplus for the non financial business sector, from 1960 to 2015. The bottom chart shows the residuals (by year and cumulative) of a regression of net investment on (lagged) Q from 1990 to 2001. Both charts show that investment has been low relative to profits and Q since the early 2000's. By 2015, the cumulative under-investment is large, around 10% of capital.

While these two stylized facts are well established, their interpretation remains controversial. There is little agreement about the causes of these evolutions, and even less about their consequences. For instance, Furman (2015) and CEA (2016) argue that the rise in concentration suggests "economic rents and barriers to competition", while Autor et al. (2017a) argue almost exactly the opposite: they think that concentration reflects "a winner take most feature" explained by the fact that "consumers have become more sensitive to price and quality due to greater product market competition." Network effects and increasing differences in the productivity of Information Technology could also

[1]The appendix shows that alternate mark-up estimates, notably those based on Barkai (2017), yield similar results.

Figure 2: Net Investment, Profits and Q-Residuals



Notes: Annual data from US Flow of Funds accounts. Net investment, net operating surplus for Non Financial Business sector; Q for Non Financial Corporate sector.

increase the efficient scale of operation of the top firms, leading to higher concentration. The key point of these later explanations is that concentration reflects an efficient increase in the scale of operation. For short, we will refer to this hypothesis as the efficient scale hypothesis (henceforth EFS).

The evolution of profits and investment could also be explained by intangible capital deepening, as discussed in Alexander and Eberly (2016). More precisely, an increase in the (intangible) capital share together with a downward bias in our traditional measures of intangible investment could lead, even in competitive markets, to an increase in profits (competitive payments for intangible services) and a decrease in (measured) investment. We will refer to this hypothesis as the intangible deepening hypothesis (henceforth INTAN). Finally, trade and globalization can explain some of the same facts (Feenstra and Weinstein, 2017). Foreign competition can lead to an increase in measured (domestic) concentration (e.g. textile industry), and a decoupling of firm value from the localization of investment. We refer to this hypothesis as the globalization hypothesis (henceforth GLOBAL).[2]

These hypotheses are not mutually exclusive. A combination of EFS, INTAN and GLOBAL is often heard in the discussion of internet giants Google, Amazon, Facebook or Apple. INTAN can also contribute to DDC since patents and other similar assets can create barriers to entry and increase the value of managing industry regulators and the challenges of litigation.

The main contribution of our paper is to propose and test two hypothesis. We first argue that the rise in concentration in most industries reflects declining domestic competition (henceforth DDC) and not EFS. We then argue that the decline in competition is (partly) responsible for the decline in investment, after controlling for INTAN and GLOBAL.

Concentration Reflects Decreasing Domestic Competition Let us start with DDC. We take into account GLOBAL by measuring separately sales, profits and investment at home and abroad, and we adjust our measures of concentration for foreign imports, following Feenstra and Weinstein (2017). The main alternative hypothesis to DDC is then EFS. We rule out EFS with three pieces of evidence. The first piece of evidence is a comparison with Europe. We consider industries with significant increases in concentration in the U.S., such as the Telecom industry, and we show that these same industries have not experienced similar increases in concentration and profit rates in Europe, even though they use the same technology and are exposed to the same foreign competition. Secondly, EFS predicts that concentration should lead to productivity gains at the industry level, as high productivity leaders expand. There is some evidence for EFS during the 1990s as the relationship between concentration and productivity was positive, but it is zero or negative in the 2000s. Thirdly, EFS predicts that leaders should increase investment and R&D in concentrating industries. We find the opposite: the relative investment of leaders is lower in concentrated industries, in physical and intangible capital. We conclude that EFS cannot be the main explanation for concentration in most industries.

[2]One could entertain other hypotheses – such as weak demand or credit constraints – but previous research has shown that they do not fit the facts. See **?** for detailed discussions and references.

DDC Causes Low Investment The second point of our paper is that DDC causes low investment. Even if we make a convincing argument that DDC explains the observed rise in concentration, it is not obvious how this should affect investment. Investment and concentration are jointly endogenous, and in models of innovation (Klette and Kortum, 2004), rents can encourage investment in innovation. The impact of competition on investment is therefore an empirical question.[3]

The first empirical challenge is to measure investment correctly and address the INTAN hypothesis. We build on Peters and Taylor (2016) and Alexander and Eberly (2016) to take into account intangible assets. We find that mismeasured intangible investment accounts for a quarter to a third of the apparent investment gap (**?**). This paper focuses on the remaining two thirds.

The second challenge for the DDC hypothesis is to establish a *causal* connection between competition and investment. The main identification issue is that firm entry and exit are endogenous. Consider an industry j where firms operate competitively under decreasing returns to scale. Suppose industry j receives the news at time t that the demand for its products will increase at some time $t+\tau$ in the future. There would be immediate entry of new firms in the industry. As a result, we would measure a decrease in concentration (or in Herfindahl indexes) followed and/or accompanied by an increase in investment. Anticipated demand (or productivity) shocks can thus explain why we see more investment in less concentrated industries even if it is not due to competition.

We construct three tests to show that DDC causes low investment, using changes in competition that are not driven by anticipated demand or supply shocks. We first consider industries exposed to Chinese competition. This is, in a sense, the exception that proves the rule. Unlike most others, these industries have experienced an overall increase in competition. Using the approach of Pierce and Schott (2016), we show that industry leaders react to exogenous changes in foreign competition by increasing their investment, in particular in R&D. This result is consistent with the recent work of Hombert and Matray (2015). Of course, foreign competition also drives out weak domestic firms, so the overall impact on domestic investment is ambiguous (marginally negative in our sample).

The Chinese natural experiment offers clean identification, but its external validity is problematic. It identifies an *increase* in competition for a particular sector and a limited set of firms, as opposed to a broad *decline* in domestic competition. The shock is only significant for half of the manufacturing sector, or about 10% of the non-financial private economy. For these reasons it is imperative to study the impact of DDC on the remaining 90% of the non-financial private economy.

Our second test relies on a model of noisy entry. Entry rates across industries depend on expected demand – the identification problem explained above – but also on noisy signals and on

[3] By contrast, the macroeconomics of imperfect competition are well understood (Rotemberg and Woodford, 1999) and other implications of DDC are straightforward: DDC predicts higher markups, higher profits, lower real wages, and a lower labor share. As Gilbert (2006) explains, the relationship between competition and investment is rather sensitive to the details of the environment, such as the extent of property rights (exclusive or not) or the nature of innovation (cost reduction versus new product). Looking at investment is also useful because it can help us distinguish the EFS and DDC hypotheses, as explained above. Finally, the welfare implications of a significant decline in the capital stock are large. For these three reasons, we argue that it is particularly important to understand the response of investment to DDC.

idiosyncratic entry opportunities. The variation in entry rates that is orthogonal to future demand and productivity is a valid instrument for competition. This "noisy" entry is usually small, which makes it difficult to identify the effect of competition. It turns out, however, that there is a major exception in the late 1990s. During that period, we document large variations in entry rates across industries that are uncorrelated with past and future sales growth, productivity growth, analysts' forecasts, and Tobin's Q. We discuss why the peculiar features of that period – especially during the second half of the 1990's with extreme equity valuation and abundant capital funding – are likely to have created more than the usual amount of randomness in entry rates (Gordon, 2005; Anderson et al., 2010; Hogendorn, 2011; Doms, 2004). Using noisy entry as an instrument for differences in concentration across industries, we find that concentration lowers investment and causes a gap between Q and investment, as predicted by the theory. Moreover, consistent with our hypothesis and our previous evidence from manufacturing, the decline in investment comes mostly from industry leaders.

The third test is based on large mergers & acquisitions (M&A). This test is important because mergers are a significant contributor to the overall increase in concentration. It also offers a different identification strategy. The likelihood of a merger is endogenous to future demand since we expect consolidation in declining industries, but the actual realization of the transaction is (partly) random. The identification assumption here is that other factors are captured by smooth trends, while M&A transactions are lumpy. We show that, conditional on current measures of concentration and expected sales growth, a discrete increase in merger-related concentration leads to a decline in investment.

Overall, using three entirely different identification strategies, and using both firm-level and industry-level data, we find that competition encourages investment, particularly by industry leaders, and particularly in intangible assets.

Related Literature. Our paper is related to several strands of literature. There is a growing literature studying trends on competition, concentration, and entry. Davis et al. (2006) find a secular decline in job flows. They also show that much of the rise in publicly traded firm volatility during the 1990's is a consequence of the boom in IPOs, both because young firms are more volatile, and because they challenge incumbents. Haltiwanger et al. (2011) find that "job creation and destruction both exhibit a downward trend over the past few decades." Decker et al. (2015) argue that, whereas in the 1980's and 1990's declining dynamism was observed in selected sectors (notably retail), the decline was observed across all sectors in the 2000's, including the traditionally high-growth information technology sector. Furman (2015) shows that "the distribution of returns to capital has grown increasingly skewed and the high returns increasingly persistent" and argues that it "potentially reflects the rising influence of economic rents and barriers to competition."[4] CEA (2016) and Grullon et al. (2016) are the first papers to extensively document the broad increases

[4] Furman (2015) also emphasizes emphasizes the weakness of corporate fixed investment and points out that low investment has coincided with high private returns to capital, implying an increase in the payout rate (dividends and shares buyback).

in profits and concentration. Grullon et al. (2016) also show that firms in concentrating industries experience positive abnormal stock returns and more profitable M&A deals. Blonigen and Pierce (2016) find that M&As are associated with increases in average markups. Dottling et al. (2017) find that concentration has increased in the U.S. while it has remained stable (or decreased) in Europe. Faccio and Zingales (2017) show that competition in the mobile telecommunication industry is heavily influenced by political factors, and that, in recent years, many countries have adopted more competition-friendly policies than the US. Autor et al. (2017a) study the link between concentration and the labor share. An important issue in the literature is the measurement of markups and excess profits. The macroeconomic literature focuses on the cyclical behavior of markups (Rotemberg and Woodford, 1999; Nekarda and Ramey, 2013). Over long horizons, however, it is difficult to separate excess profits from changes in the capital share. De-Loecker and Eeckhout (2017) estimate markups using the ratio of sales to costs of goods sold, but in the long run this ratio depends on the share of intangible expenses, and the resulting markup does not directly provide a measure of market power. Barkai (2017), on the other hand, estimates directly the required return on capital and finds a significant increase in excess profits.

The weakness of investment has been discussed in the context of weak overall growth (IMF, 2014; Furman, 2015; Hall, 2015; Fernald et al., 2017). Alexander and Eberly (2016) emphasize the role of intangible investment. **?** show that the recent weakness of investment relative to Tobin's Q is not explained by low expected productivity growth, low expected demand, or financial frictions. Consistent with our emphasis on market power, Lee et al. (2016) find that capital stopped flowing to high Q industries in the late 1990's. A large literature, surveyed by Gilbert (2006), studies the relationship between competition, innovation and investment. Comin and Philippon (2005) find that "firm volatility increases after deregulation [and] is linked to research and development spending." Aghion et al. (2009) study how foreign firm entry affects investment and innovation incentives of incumbent firms. Varela (2017) studies the feedback effects on investment from relaxing laggards' financial constraints. She finds that improving laggards' access to funding not only increases their own investment, but also pushes leaders to invest more to remain competitive. Corhay et al. (2017) study the link between (risky) markups and expected excess returns.

Last, our paper is related to the effect of foreign competition – particularly from China (see Bernard et al. (2012) for a review). Bernard et al. (2006) show that capital-intensive plants and industries are more likely to survive and grow in the wake of import competition. Bloom et al. (2015) argue that Chinese import competition leads to increased technical change within firms and a reallocation of employment towards more technologically advanced firms. Frésard and Valta (2015) find that tariff reductions lead to declines in investment in markets with competition in strategic substitutes and low costs of entry. Within-industry, they find that investment declines primarily at financially constrained firms. The decline in investment is negligible for financially stable firms and firms in markets featuring competition in strategic complements. Hombert and Matray (2015) show that R&D-intensive firms were better able to cope with Chinese competition than low-R&D firms. They explain this result based on product differentiation, using the Hoberg and Phillips

(2017) product similarity index. Autor et al. (2013); Pierce and Schott (2016); Autor et al. (2016); Feenstra et al. (2017) study the effects of Chinese import exposure on U.S. manufacturing employment. Feenstra and Weinstein (2017) estimate the impact of globalization on mark-ups, and conclude that mark-ups decreased in industries affected by foreign competition. Some of these papers find a reduction in investment for the 'average' firm, which is consistent with our results and highlights the importance of considering industry leaders and laggards separately.

The remainder of this paper is organized as follows. Section 1 discusses our dataset and shows that the investment gap is driven by industry leaders in concentrating industries. Section 2 provides evidence of declining domestic competition. Section 3 presents the tests and results used to establish causality between competition and investment. Section 4 concludes. Various Appendices provide details on the data, mark-up estimations and robustness checks.

1 Data and Stylized Facts

In this Section we summarize the data used throughout the paper, and we present two new stylized facts that are critical to understanding the dynamics of concentration and investment.

1.1 Data

We use a wide range of aggregate-, industry- and firm-level data, summarized in Table 1. We describe the treatment of intangible assets and the calculation of Herfindahls in the rest of this section. Further details on the datasets are relegated to Appendix B.

1.1.1 Intangible Assets

It is essential to account for intangible assets when measuring capital, investment and Q. It is not always possible to use exactly the same definitions in aggregate/industry datasets and in firm-level datasets.

Aggregate and Industry-level data. Aggregate and industry-level data are sourced from U.S. and European National Accounts. Since 2013, these accounts capitalize 'identifiable' intangible assets such as software, R&D, and entertainment, literary, and artistic originals. We use the corresponding measures of I and K in our analyses. When estimating Q, we follow the literature and measure the ratio of market value to the replacement cost of capital including intangibles (**?**).

Firm-level data. US firm-level data are sourced from Compustat and therefore follow GAAP. Under GAAP, firms report stock and flow measures of tangible capital in the Property, Plant and Equipment (PP&E) and Capital Expenditures (CAPX) line items. The treatment of intangible assets, however, is more nuanced. Internally created intangibles are expensed on the income statement and almost never appear on the balance sheet – these include R&D and advertising expenses,

Table 1: Summary of Key Data Sources

Data type	**Key Data fields**	**Source**	**Region**	**Granularity**
Aggregate/sector-level	I, K, OS, and Q	Flow of Funds	US	Country and Sector (NFCB, NFNCB)
Industry-level data	I, K and OS	BEA	US	~NAICS L3
		OECD STAN	EU	ISIC Rev 4
Concentration Measures	Herfindahls and Concentration Ratios	Economic Census	US	NAICS L3-L6
		Compustat	US	BEA segments
		CompNET	EU	ISIC Rev 4
		BvD Amadeus	EU	ISIC Rev 4
Firm Financials	I, K, OS,Q and other controls	Compustat NA	US	Firm
		Compustat Global	EU	Firm
China	Import Exposure	UN Comtrade	Global	HS code
	NTR Gap and import value	Peter Schott's website	US	NAICS L6
Productivity & controls	TFP & Mfg Industry Controls	NBER-CES Database	US	NAICS L6
	TFP	BLS KLEMS	US	BEA segments
Other	Analyst Forecasts	I/B/E/S	US	Firm
	Intangible Capital	Peters & Taylor	US	Firm

for example. Externally created (i.e., acquired) intangible assets are capitalized and reported in the Intangible Assets line item. These include Goodwill and Other (identifiable) Intangible Assets such as patents and software.

Peters and Taylor (2016) (PT for short) estimate firm-level intangible capital by combining estimates of internally and externally-created intangibles. For the former, they follow Corrado and Hulten (2010) in using granular investment and depreciation assumptions on the R&D and Sales, General & Administrative (SGA) line items to capitalize R&D as well as "expenditures on product design, marketing and customer support, and human capital and organizational development." For the latter, they use the balance sheet measure of externally created intangibles directly.[5] We use PT's estimates of I and K in our firm-level analyses, and report results separately for tangible, intangible and total capital where appropriate. For Q, PT advocate a measure labeled 'total Q' and defined as the ratio of market value of productive assets to tangible plus intangible capital. We deviate from this definition and instead estimate firm-level Q as the market-to-book ratio, in line with **?**. **?** compare the distribution and performance of market-to-book and 'total Q' and find that market-to-book is more stable over time and relies on fewer measurement assumptions. Nonetheless, we confirm that our results are robust to using 'total Q'.

1.1.2 Adjusted Herfindahls

Our ideal competition measure should cover the whole economy and take into account foreign competition (i.e., imports).

For Manufacturing, Feenstra and Weinstein (2017) (FW for short) construct such a measure. They use Census Herfindahls for the U.S. and import data for foreign countries. The replication files available at the author's website include Herfindahls at the country- and 4-digit Harmonized System (HS-4) level, from 1992 to 2005. We start from these Herfindahls, aggregate them and map them to BEA segments.[6] We then extend the series to cover 1990 to 2015 by regressing FW Herfindahls on Compustat Herfindahls and share of sales.[7] The detailed calculations are described

[5] Because it includes non-identifiable assets such as Goodwill, marketing and human capital, PT's measure of intangible capital is broader than that of National Accounts. It results in higher capital estimates. Our conclusions are robust to excluding Goodwill from PT's measure of intangible capital

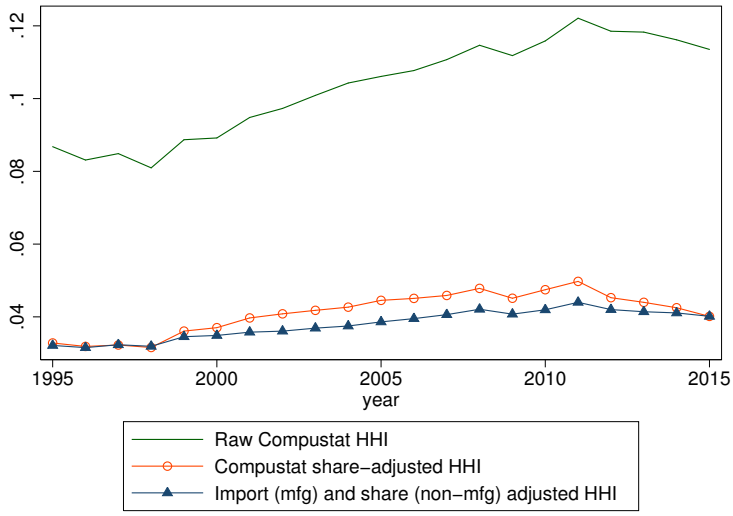
[6] First, we aggregate country-sector Herfindahls HHI_t^{cj} across countries c to obtain the *Overall Herfindahl Index* for HS-4 sector j $\left(HHI_t^j\right)$. Next, we use the correspondence of Pierce and Schott (2012) to map HS-4 sectors to NAICS-6 sectors, which can then be mapped to BEA segments (which roughly correspond to ~NAICS-3 segments). Last, we aggregate Herfindahls across HS-4 segments j into BEA segments k, to obtain Herfindahls at the BEA segment k: HHI_t^k.

[7] FW Herfindahls are based on SIC segments before 1997 and NAICS segments afterwards, which results in a jump in HHI_t^k for some series. We control for the jump by subtracting the 1997 change in HHI_t^k from all HHI_t^k series after 1998. We then extend the time series through a regression of the form $log\left(HHI_t^k\right) = log\left(HHI_t^{k,CPraw}\right) + log\left(s_{kt}^{CP}\right) + \alpha^k + \varepsilon_{kt}$, where $HHI_t^{k,CP}$ denotes the Herfindahl from Compustat and s_{kt}^{CP} denotes the share of sales of Compustat firms as a percent of total US output plus imports. The Compustat Herfindahl $\left(HHI_t^{k,CP}\right)$ is highly correlated with the FW Herfindahl $\left(HHI_t^k\right)$ at the BEA segment-level, particularly once controlling for the share of Compustat sales. For instance, the R^2 of the regression above excluding fixed effects is 42% and including fixed effects is 95% – so the filled-in Herfindahls seem accurate. The level of HHI_t^k following FW tends to be lower than the level implied by Compustat. Most of our regressions include fixed effects, so this is not an issue. However, for columns 1-2 in Table 6 as well as some Figures, the level of the HHI_t^k matters. We therefore add a constant

in the appendix.

Outside Manufacturing, neither Census nor foreign Herfindahls are available – so we have to use Compustat. We start with the "raw" Herfindahls from Compustat and adjust them to account for the domestic coverage of Compustat as well as the share of imports. Consider an industry with x firms in Compustat and N firms globally, all with equal shares of the U.S. market. The Compustat share of output is $s^{CP} = \frac{x}{N}$, and the Compustat-based Herfindahl $HHI^{CP} = \frac{1}{x}$. Under these assumptions, the adjusted Herfindahl can computed as $HHI_t^k = \frac{1}{N} = HHI_{kt}^{CP} \times s_{kt}^{CP}$ where s_{kt}^{CP} is the share of Compustat sales in US output plus imports. We refer to this measure as the "Compustat share-adjusted" Herfindahl (HHI_{kt}^{CPadj}). For service sectors, import data is not available but these are typically small, so we set them to zero.

Figure 3: Weighted Average Herfindahls



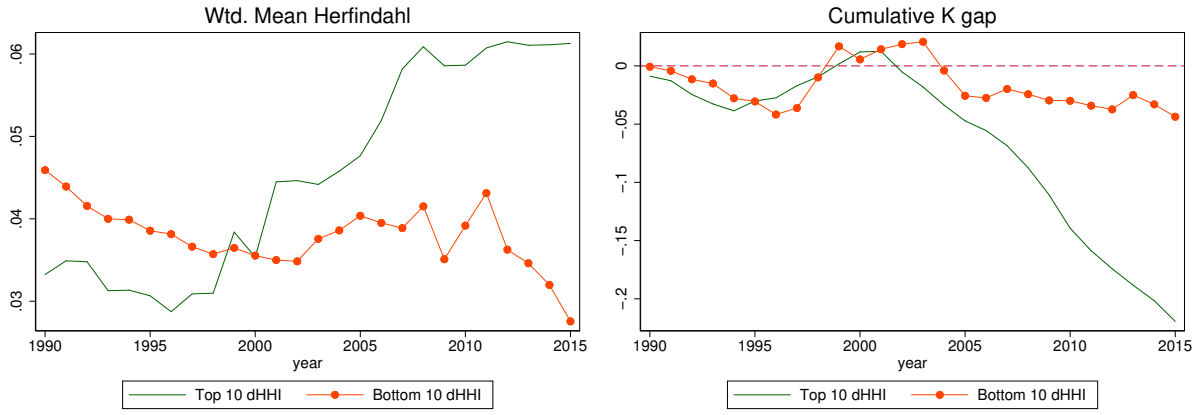
Notes: Annual data. Figure shows the weighted average of three measures of Herfindahls. The Raw Compustat HHI is the sum of squared Compustat market shares. The Compustat share-adjusted HHI adjusts for the Compustat share of sales. The Import and Share adjusted HHI is based on FW Herfindahls for Manufacturing and Compustat share-adjusted Herfindahls for non-manufacturing.

Figure 3 shows the impact of both adjustments sequentially. The Compustat share adjustment accounts for the share of Compustat sales in domestic output plus imports, while the import adjustment accounts for the concentration of foreign firms. All three series have increased since 1995, by 30%, 22% and 25%, respectively. The increase is concentrated in non-manufacturing industries as shown in Appendix B.1.4.[8]

across all manufacturing segments, to match the average level of HHI_t^k to that of $HHI_t^{k,CP}$ across all manufacturing industries.

[8]We validate the use of Compustat in two ways. First, we compare the evolution of Herfindahls adjusted for the Compustat share of sales (HHI_{kt}^{CPadj}) to alternate Compustat- and FW-based Herfindahls, as described in Appendix B.1.4. HHI_{kt}^{CPadj} exhibits the highest correlation with FW-Herfindahls (81% in levels and 66% in changes). Second, we gather census CRs and use them to (i) test the robustness of key results to using Census CRs instead of import-adjusted Herfindahls; and (ii) compare Compustat CRs against Census CRs. Most of our results are robust to using

Figure 4: Cumulative Capital Gap for Concentrating and Non-Concentrating Industries



Notes: Annual data. Left plot shows the weighted average import adjusted Herfindahl for the 10 industries with the largest and smallest relative change in import-adjusted Herfindahl. Right plot shows the cumulative implied capital gap (as percent of capital stock) for the corresponding industries. See text for details.

1.2 Two Stylized Facts

This section shows why it is critical to understand the dynamics of concentrating industries, and within industries, of the leading firms.

Fact 1: The Investment Gap Comes from Concentrating Industries. Figure 4 shows that the capital gap is coming from concentrating industries.[9] The solid (dotted) line plots the implied capital gap relative to Q for the top (bottom) 10 concentrating industries. For each group, the capital gap is calculated based on the cumulative residuals of separate industry-level regressions of net industry investment from the BEA on our measure of (lagged) industry Q from Compustat.[10] The Herfindahl index for the bottom 10 turns out to be rather stable over time, and investment remains largely in line with Q for this group.

Fact 2: Industry Leaders Account for the Increased Profit Margins and for the Investment Gap. In Table 2 (see also Appendix Figure 21), we define leaders by constant shares of

Census CRs instead of import-adjusted Herfindahls (see Appendix C for details). In addition, Census and Compustat CRs are strongly correlated at the BEA segment-level (80% in levels and 56% in changes). We also perform extensive sensitivity analyses to adjustments in the calculation of import-adjusted Herfindahls (e.g., using s_{kt}^{BEA} instead of s_{kt}^{CP}). Appendix B.1.4 provides additional details on the tests and comparisons. See Davis et al. (2006) for additional discussion of the limitations in using Compustat to measure industry concentration

[9] We define concentrating industries based on the relative change in import adjusted Herfindahls from 2000 to 2015. The top 10 concentrating industries include Arts, Health other, Inf. motion, Inf. publish and software, Inf Telecom, Transp pipeline, Transp truck, Min exOil, Retail trade, Transp_air. We exclude Agriculture because Compustat provides limited coverage for this industry.

[10] To be specific, each line is computed as follows: we first compute the residuals from separate industry-level regressions of net investment on (lagged) mean industry Q, from 1990 to 2001. Then, we average yearly residuals across the industries with the ten largest and ten smallest relative changes in import-adjusted Herfindahls from 2000 to 2015. Last, we compute the cumulative capital gap by adding residuals from 1990 to 2015, accounting for depreciation.

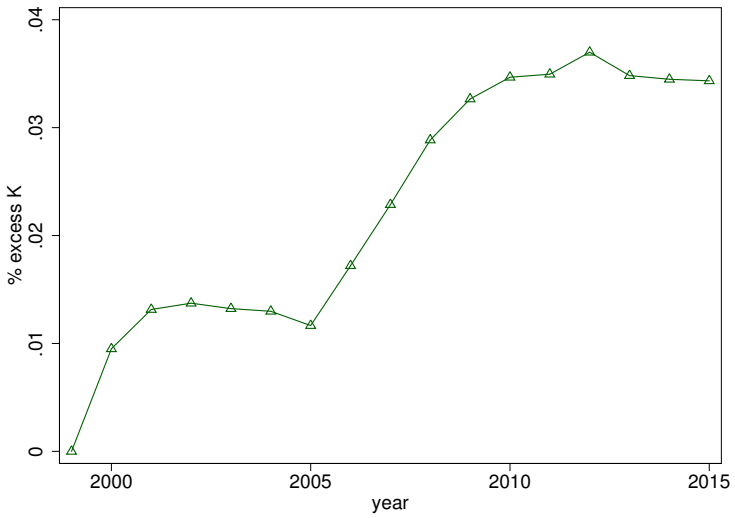
Table 2: Investment, Capital and Profits by Leaders and Laggards

Table shows the average value of a broad set of investment, capital and profitability measures by time period and market value. Leaders (laggards) include the firms with the highest (lowest) MV that combined account for 33% of MV within each industry and year. Annual data from Compustat. Lerner Index defined as $(OIBDP - DP)/SALE$.

	1980-1995			**1996-2015**			**Difference**		
	Leaders	**Mid**	**Laggards**	**Leaders**	**Mid**	**Laggards**	**Leaders**	**Mid**	**Laggards**
	0-33 pct	33-66 pct	66-100 pct	0-33 pct	33-66 pct	66-100 pct	0-33 pct	33-66 pct	66-100 pct
Share of OIBDP	0.33	0.34	0.33	0.33	0.33	0.33	**0.00**	0.00	0.00
Share of CAPX + R&D	0.35	0.32	0.34	0.29	0.32	0.39	**-0.06**	0.01	0.05
Share of PP&E	0.32	0.33	0.34	0.30	0.31	0.39	**-0.02**	-0.02	0.05
Share of K	0.31	0.33	0.36	0.27	0.33	0.39	**-0.03**	0.00	0.04
(CAPX+R&D)/OIBDP	0.71	0.67	0.71	0.52	0.57	0.70	**-0.20**	-0.10	-0.01
Lerner Index	0.10	0.09	0.07	0.13	0.11	0.07	**0.02**	0.02	0.001

market value to ensure comparability over time.[11] Capital K includes intangible capital as estimated by Peters and Taylor (2016). Table 2 shows that the leaders' share of investment and capital has decreased, while their profit margins have increased.

Figure 5: Implied Gap in K due to Leader Under-Investment



Notes: Annual data. Figure shows the cumulative implied excess capital (as percent of total U.S. capital stock for the industries in our sample) assuming Compustat leaders continue to account for 35% of CAPX and R&D investment from 2000 onward. Non-leaders assumed to maintain their observed invest levels. Excess investment assumed to depreciate at the US-wide depreciation rate. US-wide capital and depreciation data from BEA.

Table 2 suggests that leaders are responsible for most of the decline in investment relative to profits. To quantify the implied capital gap, Figure 5 plots the percentage increase in the capital stock of the U.S. non-financial private sector assuming that Compustat leaders continued to invest 35% of CAPX plus R&D from 2000 onward, while the remaining groups invested as observed. The capital stock would be ∼3.5% higher under the counter-factual. This is a large increase considering that our Compustat sample accounts for about half of investment (see Appendix B for details) and that the average annual net investment rate for the U.S. Non Financial Business sector has been less than 2% since 2002. A macroeconomic simulation by Jones and Philippon (2016) (taking into account general equilibrium effects and monetary policy) based on our implied markup series suggests a shortfall of 5 to 10%.

2 Rising Concentration Reflects Decreasing Domestic Competition

In this section we make the case that the increase in concentration reflects DDC. As explained above, we adjust our concentration measures to take into account foreign competition. The main alternative explanation is then EFS. The efficient scale argument is that technological change –

[11] OIBDP shares are stable which is consistent with stable shares of market value and stable relative discount factors. Because firms are discrete, the actual share of market value in each grouping varies from year to year. To improve comparability, we scale measured shares as if they each contained 33% of market value.

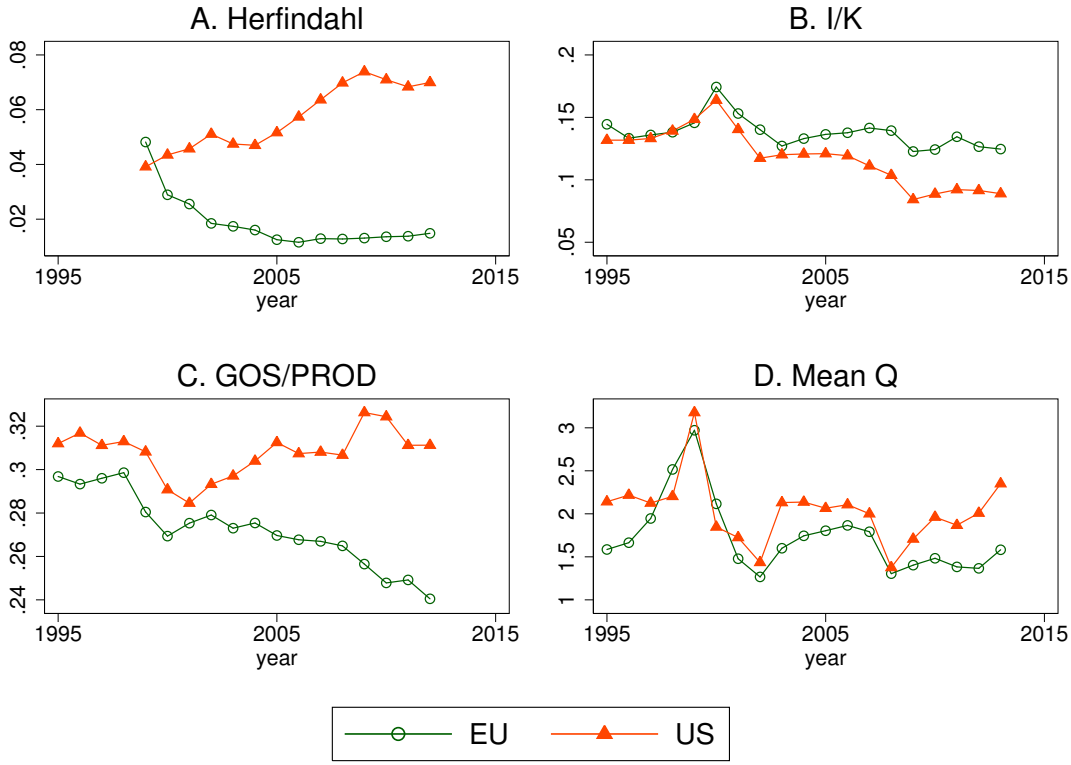
information technology, networks, winner-take-all, etc – has increased the efficient relative size of the best firms in each industry. The key point here is that increasing skewness is a efficient response to changes in the environment. We present three pieces of evidence that are inconsistent with this interpretation but consistent with DDC.

2.1 US vs. Europe

The comparison with Europe is extremely data-intensive. We rely on the dataset of Dottling et al. (2017), which includes industry- and firm-level series of profit, investment and concentration for the U.S. and Europe under consistent industry segments.[12] We present only key comparisons of industries with significant increases in concentration in the U.S. (such as Telecom). Figure 6 compares the weighted average (domestic) Herfindahl, investment rate, operating margin and Q for the 5 industries that concentrate the most in the US. We exclude the Manufacturing - Textiles industry even though it exhibits a rise in domestic concentration because the increase is primarily due to foreign competition. Accounting for imports, the Herfindahl increased much less than for the remaining 5 concentrating industries.

[12] Firm-level data is based on Compustat (NA and Global). Industry-data is based on the BEA, EU KLEMS and OECD STAN. Concentration measures are based on Compustat NA for the U.S. and BvD Orbis for Europe (given the larger presence of private firms in Europe). We are grateful to Sebnem Kalemli-Ozcan and Carolina Villegas-Sanchez for providing us with a historical time series of Herfindahls and Top-firm Market Shares computed based on the BvD Orbis merged vintage dataset of Kalemli-Ozcan et al. (2015). See Dottling et al. (2017) and Appendix B for additional details.

Figure 6: Comparison with EU for Top 5 Concentrating Industries in US



Notes: Figure based on the top 5 concentrating industries in the US. These industries are Information Telecom, Arts and Recreation, Wholesale and Retail trade, Other Services and Information Publishing (which includes software). Panel A plots the weighted average Herfindahl across these industries, weighted by sale. For the EU, each industry's Herfindahl is the weighted average Herfindahl across countries. Panel B plots the weighted average investment rate, weighted by the capital stock. Panel C plots the the weighted average ratio of Gross Operating Surplus to Production. Last, Panel D plots the weighted average mean Q, by assets. All weights are based on the U.S. share of industries to control for differences in industry sizes across regions.

The series are aggregated across industries based on US share of sales, capital, output and assets (respectively) to ensure a common weighting across regions.[13] Concentration, profits and Q increased in the U.S., while investment decreased. By contrast, concentration decreased in Europe, and investment remained (relatively) stable despite lower profits and lower Q. This true even though these industries use the same technology and are exposed to the same foreign competition. As shown in the Appendix C.1.1, these conclusions remain when looking at the underlying industries – such as Telecom and Airlines.[14] EFS, GLOBAL and INTAN therefore cannot explain the concentration in the US. On the other hand, these trends are consistent with DDC since antitrust enforcement

[13] We present results using BvD Orbis Herfindahls, and also confirm that conclusions are robust to using Concentration Measures from the ECB's CompNET (see Appendix C.1.1 for details).

[14] Airlines is not included in Figure 6 because EU KLEMS combines the entire Transportation and Storage sector, hence was combined in the analyses of Dottling et al. (2017). But we can compare concentration and mark-up trends using the ECB's CompNET.

in Telecom and Airlines has indeed become more aggressive in Europe than in the US in recent years (see Faccio and Zingales (2017) for Telecoms, Economist (2017) for Airlines, and **?** for all industries).

2.2 Concentration and TFP

According to the EFS hypothesis, concentration reflects an *efficient* increase in the scale of operation. A key prediction of the EFS hypothesis is therefore that concentration leads to productivity gains at the industry level, as high productivity leaders expand. It has happened before, for instance in Retail Trade during the 1990's.[15] The question is whether EFS is the main driver of concentration over the past 20 years as hypothesized by Autor et al. (2017a). To test this idea, we study the relationship between changes in concentration and changes in industry TFP at two levels of granularity. First, we study the more granular NAICS Level 6 manufacturing industries using productivity measures from the 2017 release of the NBER-CES database (which contains data up to 2011). Next, we broaden the sample to all US industries by using KLEMS, at the expense of considering more aggregated ∼NAICS Level 3 industries.[16] For all analyses, we consider domestic concentration to align with domestic TFP estimates. We only include industry segments that remain stable over each 5-year period in our regressions, so that no aggregation/mapping is necessary.

Table 3: Industry regressions: Concentration vs. TFP

Table shows the results of industry-level OLS regressions of contemporaneous changes in TFP and Concentration over the periods specified. Observations are weighted by value added. Columns 1-3 include NAICS-6 manufacturing industries, with TFP from NBER-CES database. Columns 4-5 include all industries in our sample, with TFP from U.S. KLEMS. Standard errors in brackets. + p<0.10, * p<0.05, ** p<.01. † TFP change to 2011 in column 3, and to 2014 in the last 5Y period of column 5 due to data availability.

	(1)	(2)	(3)	(4)	(5)
	ΔTFP$(t, t-5)$			ΔTFP$(t, t-5)$	
	97-02	02-07	07-12†	90-00	00-14†
ΔCensus CR8$(t, t-5)$	**1.456****	0.237	-1.35		
	[0.312]	[0.652]	[0.871]		
ΔCP CR8$(t, t-5)$				**0.461***	**-0.208**+
				[0.198]	**[0.115]**
Sectors	Mfg			All	
Granularity	NAICS-6			KLEMS	
Observations	469	469	299	86	129
R^2	0.045	0	0.008	0.061	0.025

Table 3 reports the results. Columns (1) to (3) focus on NAICS Level 6 manufacturing industries. As shown, the relationship is positive and significant over the 1997 to 2002 period but not after.

[15] The Retail Trade industry became substantially more concentrated – and more productive – over the 1990's. Lewis et al. (2001) find that over the 1995 to 2000 period, a quarter of the U.S. productivity growth is attributable to advances in the retail industry, and almost a sixth of that is attributable to Walmart.

[16] When necessary, we use the sales-weighted average to aggregate concentration ratios across NAICS Level 3 segments to match the granularity of KLEMS.

In fact, the relationship is negative in the 2007 to 2012 period.[17] Columns (4) and (5) show that the results are similar (and more significant) when we broaden the scope to all industries in our sample.[18] The positive relationship at the beginning of the sample is consistent with the results in Autor et al. (2017b), but the results in the 2000's are not. To be clear, Autor et al. (2017b) make two points. The first is that economic activity has shifted towards firms with lower labor shares, a fact also documented by Kehrig and Vincent (2017) and that we replicate in our data. The second point is that the concentration is explained by EFS. We find some evidence in favor of EFS in the 1990s, but evidence against it in the 2000s.

2.3 Investment by Leaders

According to the EFS hypothesis, leaders should increase investment in concentrating industries, reflecting their increasing relative productivity. We test this at the firm-level, by performing the following regression for firm i that belongs to BEA segment k:

$$\begin{aligned}\Delta \log(K_{it}) =& \beta_1 Q_{it-1} + \beta_2 HHI^k_{t-1} \times Leader^k_{it-1} + \beta_3 HHI^k_{t-1} \\ &+ \beta_4 Leader^k_{it-1} + \beta_5 \log(Age_{it-1}) + \eta_t + \mu_i + \varepsilon_{it},\end{aligned} \tag{1}$$

where K_{it} is firm capital (PP&E, Intangibles, or Total), HHI^k_t the import-adjusted Herfindahl, and $Leader^k_{it}$ is an indicator for a firm having a market value in the top quartile of segment k. We include Q_{it-1} and $\log(Age_{it-1})$ as controls, along with firm and year fixed effects (η_t and μ_i). β_2 is the coefficient of interest. Table 4 shows that leaders in concentrated industries under-invest. This is inconsistent with EFS and consistent with DDC. Appendix C.1.2 reports results using Census-based measures of concentration, and including the Noisy Entry instrument (defined below) instead of Herfindahls as an exogenous measure of competition. In unreported tests, we confirm that results are robust considering manufacturing and non-manufacturing industries separately.

[17] The number of observations decreases in column 3 due to substantial changes to NAICS Level 6 categories between NAICS 2007 and NAICS 2012. Results before 2007 are robust to considering only those industries with consistent segments from 1997 to 2012. In unreported tests, we find a negative and significant coefficient when considering the 10Y period from 2002 to 2012

[18] In unreported tests, we find positive correlations between concentration and value-added per worker, but this would be true under any model of increasing market power irrespective of productivity.

Table 4: Investment by Leaders

Table shows the results of firm-level panel regressions of the log change in the stock of capital (deflated to 2009 prices) on import-adjusted Herfindahls. Regression from 2000 to 2015, following equation (1). We consider three measures of capital: PP&E, intangibles defined as in Peters and Taylor (2016) and their sum (total). Leaders measured as the two-year moving average of an indicator for a firm having market value in the top quartile of the corresponding BEA segment k. Q and log-age included as controls. As shown, leaders decrease investment with concentration, rather than increase it. Annual data, primarily sourced from Compustat. Standard errors in brackets, clustered at the firm-level. + p<0.10, * p<0.05, ** p<.01.

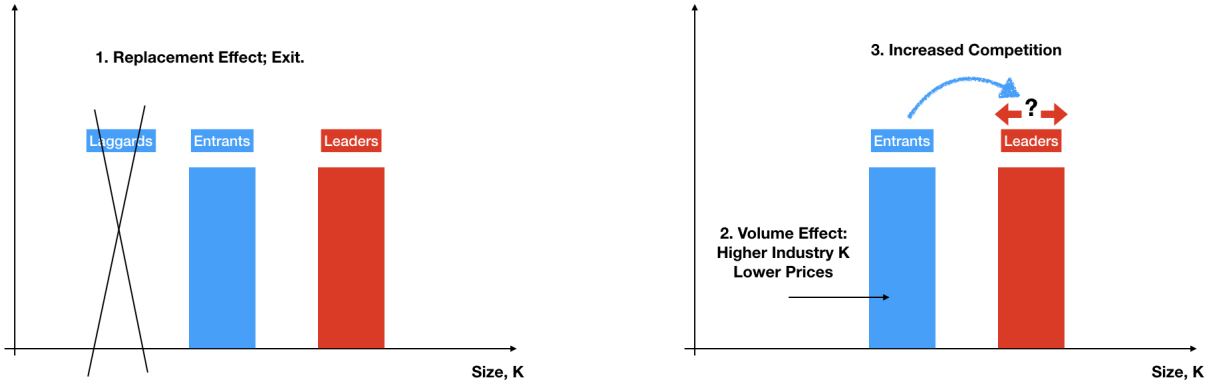
	(1) $\Delta log(PPE)^a$ ≥ 2000	(2) $\Delta log(Int_{PT})^b$ ≥ 2000	(3) $\Delta log(K_{PT})^{a+b}$ ≥ 2000
Q_{it-1}	6.84**	3.38**	4.01**
	[0.24]	[0.13]	[0.13]
HHI^k_{t-1}	15.82	11.25	21.32*
	[14.43]	[9.40]	[9.44]
$Leader^k_{it-1}$	0.91	0.37	0.22
	[1.16]	[0.96]	[0.85]
$HHI^k_{t-1} \times Leader^k_{it-1}$	**-34.41***	**-24.43+**	**-29.28****
	[13.92]	**[12.89]**	**[11.20]**
$\log(Age_{it-1})$	-6.10**	-14.02**	-12.52**
	[1.38]	[0.89]	[0.87]
Observations	59361	56472	56704
R^2	0.06	0.08	0.09
Year FE	YES	YES	YES
Firm FE	YES	YES	YES

3 Competition Encourages Investment

The previous section has shown that international, industry, and firm level evidence is inconsistent with EFS and consistent with DDC in the US. We now make the case that competition increases investment, and therefore that DDC has caused a shortfall in business investment. Establishing causality is challenging because entry, exit – and therefore concentration – are endogenous. We thus propose three different identification strategies. Figure 7 summarizes the testable predictions. Consider an industry, initially in equilibrium with some leaders and some laggards, but disrupted by entrants that are more productive than the current laggards. There is first a replacement effect, as the laggards are forced out. Then, because the entrants are productive, industry output expands and prices fall. Finally, the leaders react. This third effect is theoretically ambiguous, as discussed at length in the literature (Gilbert, 2006). In non-strategic models (Klette and Kortum 2004, monopolistic competition with iso-elastic demand curves, etc.), leaders would cut investment. In strategic models (entry deterrence, neck-and-neck competition, etc.) leaders could increase investment and innovation.

Which of these predictions we can test depends on the context. If competition is domestic, we can test the industry level response of investment, as well as the response of leaders. If entrants

Figure 7: Testable Predictions



are foreign competitors we can only test the investment response of the leaders, because there is no reason to expect domestic investment to increase.

3.1 Evidence from Chinese Competition

Our first test is based on increased competition from China during the 2000's, following Autor et al. (2016) and Pierce and Schott (2016). Pierce and Schott (2016) exploit changes in barriers to trade following the United States granting Permanent Normal Trade Relations (PNTR) to China. PNTR became effective on December 2001 as China entered the WTO.

Chinese competition leads to a strong replacement effect, consistent with Figure 7. Figure 8 shows the normalized number of firms in industries with high and low Chinese import penetration. Both groups have the same pre-existing trends, including during the dot-com boom, but start to diverge after 2000. The number of firms in industries with high import penetration decrease much faster than the number of firms in industries with low import penetration.

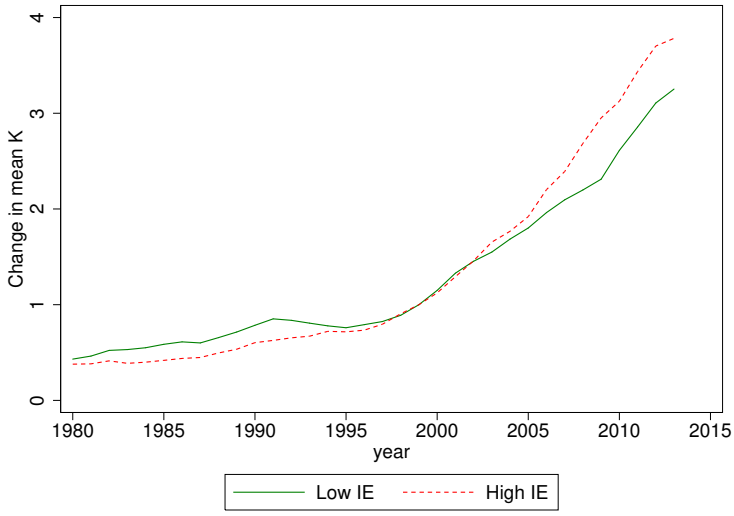
Figure 8: Number of firms by Chinese exposure (1991 = 1)



Notes: Annual data. Firm data from Compustat; import data from UN Comtrade. Manufacturing industries only, split into 'high' (above-median) and 'low' (below-median) exposure based on import penetration from 1991 to 2011.

Let us now focus on the surviving firms. Figure 9 plots the average stock of K across Compustat firms in a given year, split by the level of import exposure. As shown, average K increased faster in high exposure industries than low exposure industries. Moreover, the increase within high exposure industries is concentrated in leaders (Figure 27 in the Appendix).

Figure 9: Change in average firm K^{PT} by Chinese Exposure (1999 = 1)



Notes: Annual data. US incorporated firms in manufacturing industries only. Industries assigned to high (low) exposure if they have above (below) median NTR gap (see below for definition). Similar patterns for PP&E and Intangibles.

Figure 7 above is based on actual import penetration and one can worry about endogeneity. In our regressions (and graphical analyses of K such as Figure 9) we therefore use the instrument proposed by Pierce and Schott (2016). Before PNTR, China was considered a non-market economy which, under the Smoot-Hawley Tariff Act of 1930, are subject to relatively high tariff rates (known as "Non-Normal Trade Relations" tariffs or "non-NTR rates"). From 1980 onward, U.S. Presidents began temporarily granting NTR tariff rates to China, but required annual re-approval by congress. The re-approval process introduced substantial uncertainty around future tariff rates and limited investment by both U.S. and Chinese firms (see Pierce and Schott (2016) for a wide range of anecdotal and news-based evidence). This ended in 2000, when the U.S. granted PNTR to China. The granting of PNTR removed uncertainty around tariffs, leading to an increase in competition. Pierce and Schott (2016) show that industries facing a larger NTR gap experienced a larger increase in Chinese imports and a larger decrease in U.S. employment. We quantify the impact of granting PNTR on industry j as the difference between the non-NTR rate (to which tariffs would have risen if annual renewal had failed) and the NTR tariff rate that was locked in by PNTR

$$NTR\,Gap_j = Non\,NTR\,Rate_j - NTR\,Rate_j.$$

This measure is plausibly exogenous to industry demand and technology after 2001. The vast majority of the variation in NTR gaps is due to variation in non-NTR rates set 70 years prior to passage of PNTR. See Pierce and Schott (2016) for additional discussion. We then examine the link between increased competition and investment (by leaders and laggards) using a generalized difference-in-differences (DiD) specification:

$$\begin{aligned}\log(K_{i,j,t}) = {} & \beta_1 Post-2001 \times NTR\,Gap_j \times \overline{\Delta IP_t^{US}} \\ & + \beta_2 Post-2001 \times NTR\,Gap_j \times \overline{\Delta IP_t^{US}} \times Leader_{i,j,0} \\ & + \boldsymbol{Post-2001} \times \boldsymbol{X_{j,91}}'\boldsymbol{\gamma} + \eta_t + \mu_i + \varepsilon_{it},\end{aligned} \tag{2}$$

where the dependent variable is a given measure of capital for firm i in industry j during year t. $\overline{\Delta IP_t^{US}}$ captures time-series variation in Chinese competition averaged across all industries.[19] The first two terms on the right-hand side are the DiD terms of interest. The first one is an interaction between the NTR gap and $\overline{\Delta IP_t^{US}}$ for the post-2001 period. The second term adds an indicator for leader firms to capture differences in investment between leaders and laggards. The third term interacts the post-PNTR dummy with time-invariant industry characteristics such as initial capital and skill intensity.[20] We include year and firm fixed effects η_t and μ_i. Our main sample for this

[19] The appendix presents results excluding $\overline{\Delta IP_{j,t}^{US}}$ to mirror the specification of Pierce and Schott (2016), as well as following the approach of Autor et al. (2016) – which instruments $\Delta IP_{j,t}^{US}$ with the import penetration of 8 other advanced economies ($\Delta IP_{j,t}^{OC}$). $\Delta IP_{j,t}^{US}$ is defined in the Appendix, following Autor et al. (2016).

[20] These industry characteristics are sourced from the NBER-CES database. They include initial year (1991) (i)

analysis includes all U.S. incorporated manufacturing firms in Compustat over the 1991 to 2015 period, but we also report results only with continuing firms (i.e., firms that were in the sample before 1995 and after 2009).

Table 5 shows that leaders increase investment in response to exogenous changes in foreign competition. We consider three different measures of capital: PP&E, Intangibles (measured as in PT) and total capital (equal to the sum of PP&E and Intangibles).[21] This supports a strategic interaction/neck-to-neck competition model, where leaders invest more to deter entry, while laggards reduce investment or exit. Columns 4 to 6 focus on continuing firms; and show that leaders invested more than laggards, even when compared only to firms that survived the China shock.

Our results are consistent with Frésard and Valta (2015) and Hombert and Matray (2015). Frésard and Valta (2015) find a negative average impact of foreign competition in industries with low entry costs and strategic substitutes. They briefly study within-industry variation, and find that investment declines primarily at financially constrained firms. Hombert and Matray (2015) studies within-industry variation with a focus on firm-level R&D intensity. They show that R&D-intensive firms exhibit higher sales growth, profitability, and capital expenditures than low-R&D firms when faced with Chinese competition, consistent with our finding of increased intangible investment. They find evidence of product differentiation using the index of Hoberg and Phillips (2017). In Appendix C.2.1 we study the dynamics of employment and find that leaders increase both capital and employment, while laggards decrease both. Employment decreases faster than capital so that K/Emp increases in both groups of firms.

3.2 Evidence from Noisy Entry

The China shock provides clean identification, but it does not have clear external validity for the entire US economy. It identifies an *increase* in competition for a particular sector (manufacturing) and a limited set of firms which account for about 10% of the non-financial private economy. This section presents our second test, which broadens the sample to the entire non-financial private economy, while considering both *increases* and *decreases* in competition.

Our identification is based on the idea of noisy entry. Appendix D presents a formal model that can be summarized as follows

$$\begin{aligned} {}^{I_t}/_{K_t} &= F\left(D_t, N_t\right) \\ N_t &= \left(1-\delta\right) N_{t-1} + \gamma\left(D_t + u_t\right) + \epsilon_t \end{aligned}$$

where ${}^{I_t}/_{K_t}$ is the investment rate, D_t is industry demand, N_t is the number of firms active at t, the shocks u_t and ϵ_t are uncorrelated with D_t and the function F is increasing in both arguments. The impact of competition on investment is measured by ${}^{\partial F}/_{\partial N}$. Strategic entry $G\left(D_t + u_t\right)$ depends

percent of production workers, (ii) ratio of capital to employment; (iii) ratio of capital to value added; (iv) average wage; (v) average production wage; and (vi) an indicator for advanced technology industries.

[21] In unreported robustness tests, we confirm that our results are robust to including only balance sheet intangibles or excluding goodwill in the PT measure.

Table 5: Chinese Competition: $\log(K_t)$ results based on $NTRGap_j \times \overline{\Delta IP^{US}_{j,t}}$

Table shows the results of firm-level panel regressions of measures of capital on $NTRGap_j \times \overline{\Delta IP^{US}_{j,t}}$, following equation (2). We consider three measures of capital: PP&E, intangibles defined as in Peters and Taylor (2016) and their sum (total). Regression over 1991 - 2015 period. Leaders defined as firms with MV in top quartile as of 1999 within each NAICS Level 6 industry. Industry controls include measures of industry-level production structure (e.g., PPE/Emp). Annual data, primarily sourced from Compustat. Only US-incorporated firms in manufacturing industries included. Standard errors in brackets, clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01.

	(1) $log(PPE_t)^a$	(2) $log(Int_t^{PT})^b$	(3) $log(k_t^{PT})^{a+b}$	(4) $log(PPE_t)^a$	(5) $log(Int_t^{PT})^b$	(6) $log(k_t^{PT})^{a+b}$
$Post01 \times NTRGap$	-7.136**	-1.096	-1.223	-6.901*	-2.236	-2.075
	[2.56]	[1.99]	[1.61]	[2.75]	[1.87]	[1.56]
$Post01 \times NTRGap \times Lead_{99}$	**7.251****	**6.143****	**5.795****	**5.848***	**7.097****	**6.469****
	[2.22]	**[1.31]**	**[1.33]**	**[2.31]**	**[1.58]**	**[1.55]**
Observations	29854	29980	29982	13988	14009	14021
R^2	0.088	0.508	0.46	0.131	0.541	0.496
Industry controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Sample	All firms			Continuing firms		

on a noisy signal of (future) demand and ϵ_t captures random changes in entry costs.[22] The second equation makes it clear that running an OLS regression of investment on the number of firms (or any other measure of concentration) leads to (upwardly) biased estimates. On the other hand, both u and ϵ would be valid instruments for N.

Measuring Noisy Entry Noisy signals and idiosyncratic entry opportunities represent a temporary shock (too much or too little realized entry) that dissipates over time, generating an impulse response structure. Consistent with overall efficiency, noisy entry is usually small, and realized entry is typically consistent with (past and future) sales and productivity growth. However, there is a major exception in the late 1990s. During that period, we find large residuals in realized entry rates, controlling for fundamentals. In particular, we let noisy entry during the 1990's be the residuals from a regression of $\Delta logN_{j,91-00}$ on observables:

[22] There is a long literature showing that early entry is strategically important, in particular because of brand preferences. See Bronnenberg et al. (2012). Therefore firms have strong incentives to make risky entry decisions. See also Mongey (2016) for a model of cross-regional variation in market concentration. One can formally compute the impulse response of investment to entry in standard DSGE models such as Corhay et al. (2017).

$$\begin{aligned}\Delta log N_{j,91-00} = \beta_0 &+ \beta_1 Med\, Q_{j,91-00} + \beta_2 Med\, \Delta log\, Sales_{j,91-00} \\ &+ \beta_3 OS/K_{j,91-00} + \beta_4 CF/Assets_{j,91-00} + \beta_5 Med\, EPS\, Fcst_{j,00} \\ &+ \beta_7 \Delta IP^{US}_{j,91,99} + \beta_8 Mean\, firm\, assets_{90} + \beta_9 Mean\, firm\, age_{90} + \varepsilon_j\end{aligned}$$

where we include measures of (past and projected) profitability, sales growth, import competition, cash flow and Q, among others. The sub-index 91-00 denotes the average value from 1991 to 2000; and $Med\, EPS\, Fcst_{j,00}$ denotes the median analyst-projected long term growth in Earnings-Per-Share across all firms in industry j as of 2000.[23]

Figure 10 plots $Noisy\, Entry_{j,90-99}$ (x-axis) against the log-change in the number of firms in the 2000's (y-axis). As shown, we find large – positive and negative – variation in noisy entry across industries. Some industries, like 'Arts' and 'Accommodation' experienced substantially more entry than predicted by fundamentals. Other industries (e.g., Mining - Support) experienced too little entry. Consistent with the impulse response structure, industries that experienced more noisy entry also experienced more net exit in the 2000's.[24] These industries have lower concentration (Herfindahl) in 2000 (Appendix Figure 29). And, perhaps more importantly, noisy entry does not predict future demand or productivity. In fact, the coefficient for noisy entry shows the wrong sign (Appendix Table 18). Combined, these results suggest that our measure of noisy entry is consistent with the corresponding models; and is therefore a valid instrument for concentration.[25]

[23] All variables our are based on Compustat, except for OS/K which is based on BEA figures. This regression yields an $R^2$of 70% . We also considered absolute changes in the number of firms during the 1990's and found largely consistent results. Long term growth forecasts are often interpreted as 5-year growth forecasts.

[24] Granted, the number of firms decreased across most industries by much more than the level of noisy entry, but the decrease is more pronounced in industries with higher noisy entry. The decreasing number of firms across all industries is consistent with the aggregate trend towards concentration.

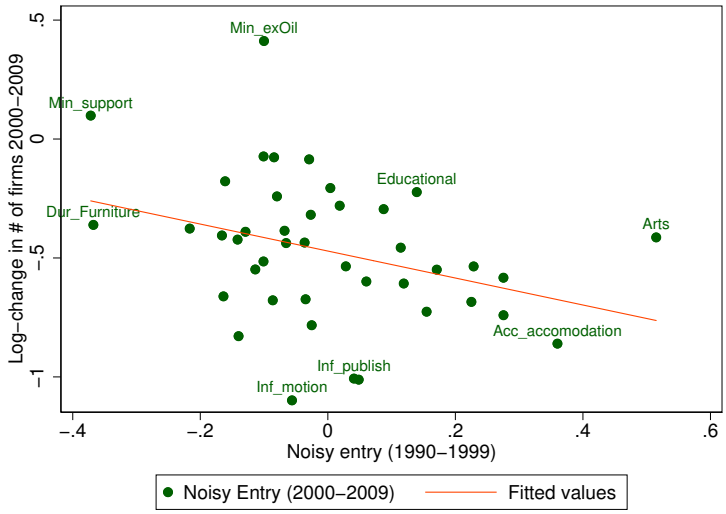
[25] The presence of noisy entry is documented for specific industries in several papers. For instance, Doms (2004) studies noisy entry and investment in the IT sector broadly – and the corresponding sub-sectors. He concludes that a "reason for the high growth rates in IT investment was that expectations were too high, especially in two sectors of the economy, telecommunications services and the dot-com sector." And Hogendorn (2011) documents excessive entry in parts of the Telecom sector.

We do not need to take a stand on whether the exuberance of the late 1990's was rational or not. Perhaps there were Bayesian mistakes, perhaps there were overly-optimistic forecasts, perhaps there were bubbles driven by the option to re-sell to future optimistic investors as in Scheinkman and Xiong (2003). All that matters for us is that these factors created variation in entry rates across industries (say in 2000) that turn out to be orthogonal to future demand (say in 2005). However, it is perhaps not surprising that we find noisy entry in the 1990s.

One explanation is potential variations in the willingness of investors (venture capitalists, or market participants in general) to fund risky ventures. This is particularly true given the optimistic environment in the late 1990's and the large inflows into Venture Capital (VC). According to the National Venture Capital Association, annual VC commitments surged during the bubble period, growing from about $10 billion in 1995 to more than $100 billion in 2000. They then receded to about $30 billion/year for the next decade (NVCA (2010)). According to Gompers and Lerner (2001), about 60 percent of VC funding in 1999 went to information technology industries, especially communications and networking, software, and information services. About 10 percent went into life sciences and medical companies, and the rest is spread over all other types of companies. Obviously, not all entry is funded by VC firms, so this can only explain a portion of the variation in entry rates – but the wide dispersion, and strong industry focus highlights the differential impact of the dot-com bubble across industries.

Another explanation is the presence of large stock market variations across most industries, as documented by Anderson et al. (2010). These extreme valuations translated into noisy entry – especially because firm entry increases

Figure 10: Change in # of firms post-2000 vs. Noisy entry pre-2000



Notes: Annual data from Compustat. See text for details on noisy entry measure.

Empirical Results We estimate the effect of competition on investment with the following industry-level panel regressions:

$$HHI_{j,t-1} = \theta_0 + \theta_1 Noisy\,Entry_{j,90-99} + \theta_2 Mean\,Q_{j,t-1} + \theta_3 Excess\,Inv_{j,90-99} + \varepsilon_{1,jt}, \quad (3)$$

and

$$\frac{NI_{jt}}{K_{jt-1}} = \beta_0 + \beta_1 \widehat{HHI_{j,t-1}} + \beta_2 Mean\,Q_{j,t-1} + \beta_3 Excess\,Inv_{j,90-99} + \varepsilon_{2,jt}. \quad (4)$$

We use noisy entry during the 1990's as an instrument for the industry-level (import adjusted) Herfindahl. We expect θ_1 to be negative because more entry leads to a lower Herfindahl. If competition (i.e., lower Herfindahl) increases investment β_1 should be negative. A potential concern with our identification strategy is that optimistic valuations may have led to excess investment among existing firms. This would bias against finding an impact of competition since investment would be lower in industries with a lot of entry. We therefore control for industry-level excess investment in the 1990's, constructed by regressing net investment on industry Q, age and size.

Equation (4) excludes industry and year fixed effects because noisy entry is constant over time. However, we can take advantage of the impulse response-structure of noisy entry to construct two additional time-varying tests. First, because noisy entry is temporary and expected to revert, it's effect on industry investment should decrease over time. We test this by studying the behavior of γ_1 in separate year-by-year regressions of net investment on noisy entry:

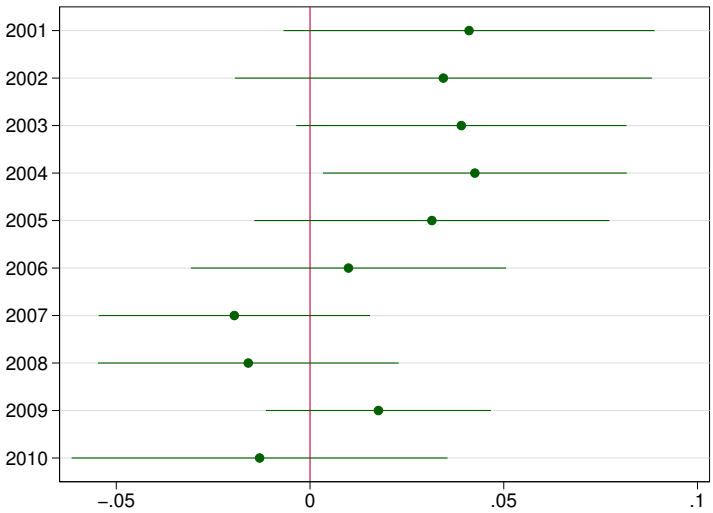
precisely during periods of high-growth such as the late 1990's (Asturias et al. (2017); Hobijn and Jovanovic (2001)).

$$\frac{NI_{j,t}}{K_{j,t-1}} = \gamma_0 + \gamma_{1t} Noisy\,Entry_{j,90-99} + \gamma_{2t} Mean\,Q_{j,t-1} + \gamma_{3t} Excess\,Inv_{j,90-99} + \varepsilon_{jt} \qquad (5)$$

Second, industries with more noisy entry in the 1990's should exhibit greater sensitivity to secular concentration trends, and therefore greater changes in I/K relative to the aggregate. We test this by interacting the sales-weighted average Herfindahl across all industries with industry-level noisy entry in equation (4). Because the aggregate Herfindahl is time-varying, we can then add industry and year fixed effects.

Figure 11 plots coefficients γ_1 from year-by-year regressions following equation (5). We include 10% confidence intervals. Consistent with the impulse response structure, Noisy Entry predicts substantially higher investment until approximately 2005 but not after. Coefficients are not always significant, but this is mostly due to the limited number of observations when running year-by-year regressions.[26]

Figure 11: Noisy entry coefficient



Notes: Figure plots the coefficient of separate year-by-year regressions of net investment on noisy entry following equation (5). Observations are weighted by the stock of capital. As shown, industries with higher noisy entry experience a temporary increase in investment. 10% confidence intervals are shown.

Table 6 reports the results of IV panel regressions following equation 4. Columns 1 and 2 show the basic regression. As expected, the coefficient on noisy entry and HHI are both negative – the first because more entry leads to a lower Herfindahl, and the second because lower competition (i.e., higher Herfindahl) leads to less investment. Columns 3 and 4 interact the weighted average Herfindahl across all industries with industry-level noisy entry. This allows us to include industry and year fixed effects. As expected, industries with more noisy entry are more sensitive to aggregate

[26] Appendix Figure 30 shows analogous results using changes in $\log(K)$ instead of I/K.

Table 6: Noisy Entry: NI/K Regression Results

Table shows the results of industry-level 2SLS regressions of net investment on Herfindahls, instrumented by noisy entry. Columns 1 and 2 focus on cross-sectional variation. Industries with higher noisy entry exhibit lower Herfindahls and higher investment. Columns 3 to 4 study time series variation and include time and industry fixed effects. They interact noisy entry with aggregate series of concentration and excess investment, and use the interactions to predict industry concentration and investment. Standard errors in brackets clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01.

	(1)	(2)	(3)	(4)
	1st St.	2nd St.	1st St.	2nd St.
	$HHI_{j,t}$	Net I/K	$HHI_{j,t}$	Net I/K
	01-05		01-15	
Mean Q (t-1)	0.02	0.025**	0.01	0.027+
	[.018]	[0.01]	[.01]	[0.01]
$Excess\,Inv_{90-99}$	-0.55	0.057		
	[1.03]	[0.45]		
$Excess\,Inv_{90-99}(i) \times NIK^{US}_{t-1}$			12.02	49.696**
			[14.33]	[16.73]
$Noisy\,Entry_{j,90-99}(i)$	**-0.15****			
	[.046]			
$Noisy\,Entry_{90-99}(i) \times Wtm\,HHI_t$			**4.41+**	
			[2.28]	
$HHI_{i,t}$		**-0.243***		**-1.278****
		[0.10]		**[0.45]**
Year FE	No		Yes	
Industry FE	No		Yes	
Observations	210		630	
RMSE	0.038		0.031	
F-stat	10.652		3.752	

concentration trends, which in turn lead to a larger reduction in investment.[27]

Appendix Table 19 performs a related analysis at the firm-level. Consistent with our hypothesis and our previous evidence from manufacturing, the increase in investment following noisy entry comes from industry leaders.

3.3 Evidence from M&A's

The third test is based on large mergers & acquisitions (M&A). As documented by prior work, merger activity is endogenous to industry dynamics. It sometimes drives consolidation in declining industries. Other times it plays an "expansionary" role following technological or regulatory shocks (Andrade and Stafford, 2004; Kaplan, 2000). Nonetheless, the actual realization of large M&A

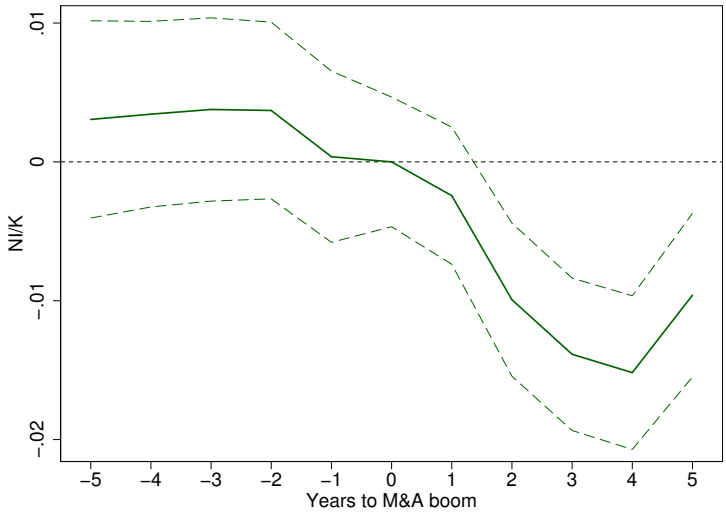
[27] Back of the envelope estimates of the implied sensitivity of investment to changes in import penetration implied by China regressions and to noisy entry suggest that they are similar. In unreported tests, we also confirm that results are robust to including only non-manufacturing industries, for which import adjustments are less material.

transactions is (partly) random. M&A typically occurs in waves, that cluster through time and across industries (Andrade et al., 2001).

Thus, we can use the discrete occurrence of M&A for identification. The identification assumption behind our test is that the omitted variables that cause the identification problem – particularly changes in demand and technology – are slow moving compared to lumpy M&A transactions. In other words, M&A waves result in sharp changes to the Herfindahl, which can identify the effect of concentration on investment without being affected by relatively slow changes in demand and technology.

We identify M&A booms as years in which firms accounting for more than 10% of sales in Compustat exit the database for M&A. This threshold selects roughly 5% of industry-year observations, mostly concentrated around M&A waves (the late 1980's, late 1990's and mid 2000's). We then study the behavior of investment around these periods. Figure 12 plots the weighted-average absolute change in the Net Investment rate over the five years before and after M&A booms, along with the corresponding 95% confidence interval. The period $t = 0$ corresponds to the fiscal year in which M&A transactions occur, so that $t = 1$ is the first complete fiscal year in which the merged firms no longer exist. As shown, investment oscillates around zero before the M&A booms, but decreases sharply thereafter. The delay in the decline is consistent with slow adjustments to investment policies; while the sharpness of the decline suggests that M&A has a discrete effect on investment, compared to (relatively) smooth changes in demand and productivity.

Figure 12: Investment Following Large M&A



Notes: Average NI/K around M&A booms, normalized by subtracting NI/K at year of M&A boom. t = 0 denotes the year of acquisitions. Observations weighted by deflated capital stock.

Table 7 confirms these results through regressions. Column 1 shows that M&A booms lead to increased (domestic) concentration. Columns 2 and 3 show that, conditional on measures of (domestic) concentration as well as expected sales growth at the time of M&A, large mergers result in lower investment. Column 2 measures M&A booms with an indicator, while column 3 considers

Table 7: M&A and NI/K

Table shows the results of industry-level OLS regressions of Net I/K on measures of M&A booms, controlling for past concentration and output growth. NI/K in percentage points. M&A boom = 1 if firms accounting for >10% of sales exit Compustat for M&A during that year. Post M&A indicator defined as years 3-5 following an M&A boom. Standard errors in brackets clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01.

	(1)	(2)	(3)	(4)
	$Herf_j^{CP}$(t)	Net I/K (All)		
	≥1980	≥1980		≥2000
$Mean\,Q_j$(t-1)		2.462**	2.468**	2.017**
		[0.66]	[0.66]	[0.48]
$Herf_j^{CP}$(t-4)	0.752**	-1.63	-1.493	
	[0.09]	[1.06]	[1.06]	
$Herf_j$(t-4)				5.707
				[4.28]
$\Delta log(Output)_j$(t-4)		5.886**	5.834**	2.709**
		[0.97]	[0.96]	[0.91]
M&A boom(t-3)	**0.025****	**-0.546+**		
	[0.01]	**[0.32]**		
Post-M&A indicator			**-0.594***	**-0.669+**
			[0.25]	**[0.33]**
Age controls	Yes	Yes		
Year FE	Yes	Yes		
Industry FE	Yes	Yes		
Observations	1530	1529	1529	688
Within R^2	0.543	0.279	0.281	0.284

years 3 to 5 following M&A booms – consistent with a relatively long-term effect of M&A on K. We use the Compustat Herfindahl – as opposed to the import-adjusted Herfindahl – to include the 1980's M&A wave;[28] and because domestic M&A deals affect domestic concentration.[29] We lag M&A booms (and concentration/sales measures) to account for the year of M&A completion and the (relatively) slow adjustment of investment policies. Column 4 shows that results are robust over the more recent period, while controlling for import adjusted Herfindahls.

4 Conclusion

US industries have become more concentrated. We argue that rising concentration in the U.S. reflects declining domestic competition (DDC) – not increasing efficient scale (EFS) – and that DDC is (partly) responsible for the low rate of investment in the U.S. Our argument for DDC rests on three pieces of evidence. First, industry-level concentration, profitability and investment trends

[28] Import-adjusted Herfindahls are only available after 1989.

[29] We cannot use census-based measures of concentration because they are available only every five years, and we are interested in sharp changes to concentration in the years surrounding M&A booms.

in Europe differ from those in the US, despite the use of similar technologies across the regions. Second, the relationship between concentration and industry productivity has been zero or negative in the 2000s. Finally, leaders invest less in physical and intangible assets in concentrated industries.

We then show that competition – actual or via the threat of entry – has a positive causal impact on investment, in particular by industry leaders. We test this idea using the well-known China Shock, as well as a model of noisy entry. We find that leaders react to exogenous increases in competition by increasing investment – and, inversely, that leaders decrease investment when competition decreases. Finally, we show that, controlling for smooth industry trends, investment decreases sharply following bursts of M&A activity.

If these conclusions are correct, they carry significant welfare implications. Decreasing competition leads to higher markups, lower real wages, and a lower labor share. In macro-economic models, the welfare losses from an investment gap driven by decreasing competition can be large. For instance, Jones and Philippon (2016) calibrate a standard DSGE model to study the macroeconomic effects of declining competition during the 2000's. They find that the capital stock is 5% to 10% lower and that the Zero Lower Bound (ZLB) binds for 2 more years than under constant competition.

References

Acemoglu, D., D. Autor, D. Dorn, G. H. Hanson, and B. Price (2016, jan). Import competition and the great US employment sag of the 2000s. *Journal of Labor Economics 34*(S1), S141–S198.

Aghion, P., R. Blundell, R. Griffith, P. Howitt, and S. Prantl (2009, feb). The effects of entry on incumbent innovation and productivity. *Review of Economics and Statistics 91*(1), 20–32.

Alexander, L. and J. Eberly (2016). Investment hollowing out. *Working Paper*.

Anderson, K., C. Brooks, and A. Katsaris (2010, jun). Speculative bubbles in the s&p 500: Was the tech bubble confined to the tech sector? *Journal of Empirical Finance 17*(3), 345–361.

Andrade, G., M. Mitchell, and E. Stafford (2001, may). New evidence and perspectives on mergers. *Journal of Economic Perspectives 15*(2), 103–120.

Andrade, G. and E. Stafford (2004, jan). Investigating the economic role of mergers. *Journal of Corporate Finance 10*(1), 1–36.

Asturias, J., S. Hur, T. Kehoe, and K. Ruhl (2017, feb). Firm entry and exit and aggregate growth. Technical report, Mimeo.

Autor, D., D. Dorn, G. H. Hanson, G. Pisano, and P. Shu (2016, December). Foreign competition and domestic innovation: Evidence from u.s. patents. Working Paper 22879, National Bureau of Economic Research.

Autor, D., D. Dorn, L. Katz, C. Patterson, and J. V. Reenen (2017a, jan). Concentrating on the fall of the labor share. Technical report, mimeo.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen (2017b, May). The fall of the labor share and the rise of superstar firms. Working Paper 23396, National Bureau of Economic Research.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review 103*(6), 2121–2168.

Autor, D. H., D. Dorn, and G. H. Hanson (2016, oct). The china shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics 8*(1), 205–240.

Barkai, S. (2017). Declining labor and capital shares. University of Chicago.

Bernard, A., J. B. Jensen, and P. K. Schott (2006). Survival of the best fit: Exposure to low-wage countries and the (uneven) growth of us manufacturing plants. *Journal of International Economics*.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2012). The empirics of firm heterogeneity and international trade. *Annual Review of Economics Vol. 4(1)*, 283–313.

Blonigen, B. A. and J. R. Pierce (2016, October). Evidence for the effects of mergers on market power and efficiency. Working Paper 22750, National Bureau of Economic Research.

Bloom, N., M. Draca, and J. Van Reenen (2015). Trade induced technical change: The impact of chinese imports on innovation, diffusion and productivity. *Review of Economic Studies*.

Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012, oct). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review 102*(6), 2472–2508.

Caballero, R. J., E. Farhi, and P.-O. Gourinchas (2017, May). Rents, technical change, and risk premia accounting for secular trends in interest rates, returns on capital, earning yields, and factor shares. *American Economic Review Papers & Proceedings 107*(5), 614–20.

CEA (2016, April). Benefits of competition and indicators of market power. Issue Brief.

Comin, D. and T. Philippon (2005). The rise in firm-level volatility: Causes and consequences. In M. Gertler and K. Rogoff (Eds.), *Macroeconomics Annual*, Cambridge. NBER.

Corhay, A., H. Kung, and L. Schmid (2017). Competition, markups and predictable returns. Working Paper.

Corrado, C. A. and C. R. Hulten (2010, May). How do you measure a "technological revolution"? *American Economic Review 100*(2), 99–104.

Davis, S., J. Haltiwanger, R. Jarmin, and J. Miranda (2006). Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. In D. Acemoglu, K. Rogoff, and M. Woodford (Eds.), *Macroeconomics Annual*, Cambridge. NBER.

De-Loecker, J. and J. Eeckhout (2017). The rise of market power and the macroeconomic implications.

De Loecker, J. and F. Warzynski (2012, May). Markups and firm-level export status. *American Economic Review 102*(6), 2437–71.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2015, November). Where has all the skewness gone? the decline in high-growth (young) firms in the u.s. Working Papers 15-43, Center for Economic Studies, U.S. Census Bureau.

Doms, M. C. (2004). The boom and bust in information technology investment. *Economic Review*, 19–34.

Dottling, R., G. Gutiérrez, and T. Philippon (2017). Is there an investment gap in advanced economies? if so, why? mimeo.

Economist, T. (2017). A lack of competition explains the flaws in american aviation.

Elzinga, K. G. and D. E. Mills (2011, May). The lerner index of monopoly power: Origins and uses. *American Economic Review 101*(3), 558–64.

Faccio, M. and L. Zingales (2017, January). Political determinants of competition in the mobile telecommunication industry. Working Paper 23041, National Bureau of Economic Research.

Feenstra, R. (2003). A homothetic utility function for monopolistic competition models, without constant price elasticity. *Economic Letters 78*, 79–86.

Feenstra, R., H. Ma, and Y. Xu (2017, March). The china syndrome: Local labor market effects of import competition in the united states: Comment. Working paper.

Feenstra, R. C. and D. E. Weinstein (2017). Globalization, markups and u.s. welfare. *Journal of Political Economy*.

Fernald, J. G., R. E. Hall, J. H. Stock, and M. W. Watson (2017, 2017). The disappointing recovery of output after 2009. *Brookings Papers on Economic Activity*.

Frésard, L. and P. Valta (2015). How does corporate investment respond to increased entry threat? *The Review of Corporate Finance Studies*.

Furman, J. (2015). Business investment in the united states: Facts, explanations, puzzles, and policies. Technical report, Council of Economic Advisers.

Gali, J. (2008). *Monetary Policy, Inflation, and the Business Cycle*. Princeton University Press.

Gilbert, R. (2006). Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate? In *Innovation Policy and the Economy, Volume 6*, NBER Chapters, pp. 159–215. National Bureau of Economic Research, Inc.

Gompers, P. and J. Lerner (2001, may). The venture capital revolution. *Journal of Economic Perspectives 15*(2), 145–168.

Gordon, R. (2005, nov). The 1920s and the 1990s in mutual reflection. Technical report.

Grullon, G., Y. Larkin, and R. Michaely (2016). Are us industries becoming more concentrated. Technical report, Working paper.

Gutiérrez, G. (2017). Investigating global labor and profit shares.

Hall, R. E. (1988). The relation between price and marginal cost in u.s. industry. *Journal of Political Economy*.

Hall, R. E. (2015, jan). Quantifying the lasting harm to the US economy from the financial crisis. *NBER Macroeconomics Annual 29*(1), 71–128.

Haltiwanger, J., R. Jarmin, and J. Miranda (2011, March). Historically large decline in job creation from startup and existing firms in the 2008–2009 recession. Census Bureau Report.

Hoberg, G. and G. Phillips (2017). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*.

Hobijn, B. and B. Jovanovic (2001, dec). The information-technology revolution and the stock market: Evidence. *American Economic Review 91*(5), 1203–1220.

Hogendorn, C. (2011, dec). Excessive(?) entry of national telecom networks, 1990–2001. *Telecommunications Policy 35*(11), 920–932.

Hombert, J. and A. Matray (2015). Can innovation help u.s. manufacturing firms escape import competition from china? Technical report, HEC Paris Research Paper No. FIN-2015-1075.

IMF (2014). Private investment: What's the holdup? *Chapter 4, World Economic Outlook*.

Imrohoroglu, A. and S. Tuzel (2014). Firm-level productivity, risk, and return. *Management Science*.

Jones, C. and T. Philippon (2016). The secular stagnation of investment?

Kalemli-Ozcan, S., B. Sorensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2015). How to construct nationally representative firm level data from the orbis global database. Technical report, National Bureau of Economic Research.

Kaplan, S. N. (2000). *Introduction to Mergers and Productivity.*

Kehrig, M. and N. Vincent (2017). Growing productivity without growing wages: The micro-level anatomy of the aggregate labor share decline. *Unpublished*.

Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of Political Economy 112*(5), 986–1018.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*.

Lee, D. W., H.-H. Shin, and R. M. Stulz (2016, September). Why does capital no longer flow more to the industries with the best growth opportunities? *Fisher College of Business Working Paper 2016*(15).

Lewis, B., A. Augereau, M. Cho, B. Johnson, B. Neiman, G. Olazabal, M. Sandler, S. Schrauf, K. Stange, A. Tilton, E. Xin, B. Regout, A. Webb, M. Nevens, L. Mendonca, V. Palmade, G. Hughes, and J. Manyika (2001). Us productivity growth, 1995 - 2000.

Mongey, S. (2016). Market structure and monetary non-neutrality. *Working paper*.

Nekarda, C. and V. Ramey (2013, jun). The cyclical behavior of the price-cost markup. Technical report, NBER Working paper.

NVCA (2010). The economic importance of venture capital backed companies to the u.s. economy. Technical report, National Venture Capital Association, Thomson Reuters.

Olley, G. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*.

Peters, R. H. and L. A. Taylor (2016). Intangible capital and the investment-q relation. *Journal of Financial Economics*.

Pierce, J. R. and P. K. Schott (2012). A concordance between ten-digit u.s. harmonized system codes and sic/naics product classes and industries.

Pierce, J. R. and P. K. Schott (2016, jul). The surprisingly swift decline of US manufacturing employment. *American Economic Review 106*(7), 1632–1662.

Pryor, F. L. (2001, May). New trends in u.s. industrial concentration. *Review of Industrial Organization 18*(3), 301–326.

Rotemberg, J. and M. Woodford (1999). The cyclical behavior of prices and costs. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1B. Elsevier Science, North Holland.

Scheinkman, J. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy 111*, 1183–1219.

Varela, L. (2017). Reallocation, competition and productivity: Evidence from a financial liberalization episode. Working Papers 2015-042-23, Department of Economics, University of Houston.

Appendices

Appendices A to C provide additional discussion of (i) mark-up estimation, (ii) data sources and definitions and (iii) results, respectively. Appendix D presents a model of competition and discusses its implications for investment. Because the discussion of concentration often veers towards Google, Apple, Facebook, Amazon and Microsoft, an additional Appendix (available upon request) discusses the evolution of investment and mark-up patterns at these firms.

A Estimating Mark-ups

We can further study the competitive trends in the US economy by analyzing the evolution of mark-ups. Mark-ups complement concentration measures by providing a direct measure of a firm's ability to extract rents from the market – independent of geographic/product market definitions. Such estimates remain powerful even under complex competitive structures.

We consider three different approaches – from purely empirical measures of profit margins such as the Lerner Index, to more complex estimates such as those of De-Loecker and Eeckhout (2017) (DLE for short). Despite relying on fairly different assumptions, all three approaches suggest that mark-ups increased – particularly after 2000. These results provide fairly conclusive evidence that market power (and profits) have increased.[30] Grullon et al. (2016) and Gutiérrez (2017) present evidence linking mark-ups and concentration.

A.1 Approaches

We consider three approaches:

1. **Lerner Index**: A simple, purely empirical measure of mark-ups is the operating profit rate, which we call the "Lerner index" for short.[31] We use this measure in Section 1. Grullon et al. (2016) shows that the Lerner index has increased at industries that have become more concentrated. The Lerner index carries several limitations as a measure of market power as discussed in Elzinga and Mills (2011). One issue in particular is that it does not account for changing cost of capital.

2. **User Cost-implied mark-ups:** Second, Barkai (2017) estimates the user cost of capital from expected returns on bond or equity for the U.S. Non Financial Corporate Sector. Caballero et al. (2017) implement a similar approach but without specifying a cost of capital. Gutiérrez (2017) imputes mark-ups at the industry-level from output and capital series using Equity Risk Premia. The main advantage of this approach is that it estimates excess profit directly. However, it relies on estimates for the stock of capital and the cost-of-capital.

[30] The magnitude and timing of the increases in mark-ups differs substantially across estimates, however. Such differences should be considered when selecting a preferred measure of mark-ups.

[31] The Lerner Index is defined as the ratio of operating income before depreciation minus depreciation to sales. We convert the Lerner index to a measure of mark-ups as $\mu^{LI} = \frac{1}{1-LI}$.

3. **Firm-level estimates from production data**: Last, DLE estimate firm-level mark-ups following the approach of De Loecker and Warzynski (2012), which in turn builds on Hall (1988) and Olley and Pakes (1996) (among others). It relies on cost minimization and there being (at least) one variable input of production free from adjustment cost; for which the wedge between that input's revenue share and its output elasticity is a direct measure of the firm's markup. It assumes that firms optimize against the variable input every period, and therefore requires a sizable expenditure on that item.

   DLE assume constant coefficients and use a Cobb-Douglas production function with constant coefficients.[32] They treat cost-of-goods sold (COGS) as the variable input and consider gross PP&E as the measure of capital. Using a Cobb-Douglas production function, this implies that mark-ups are a scaled series of the SALE/COGS ratio, which introduces substantial limitations. The SALE/COGS ratio mechanically increases with the share of intangible assets and expenses - so this approach sheds little light on whether market power has increased. As an alternative approach, we also estimate DLE-style mark-ups using 'total expenses' instead of COGS, defined as in Imrohoroglu and Tuzel (2014).[33]

A.2 Results

Figure 13 shows the evolution of the weighted average Lerner Index, User Cost-implied mark-ups and DLE mark-ups (with COGS and Total expenses). All estimates are based on the Compustat sample in Gutiérrez (2017) (which excludes Financials, Real Estate and Utilities).[34] The top chart shows the full time-series, while the bottom chart normalizes all series to 1 as of 1980 to focus on relative changes.

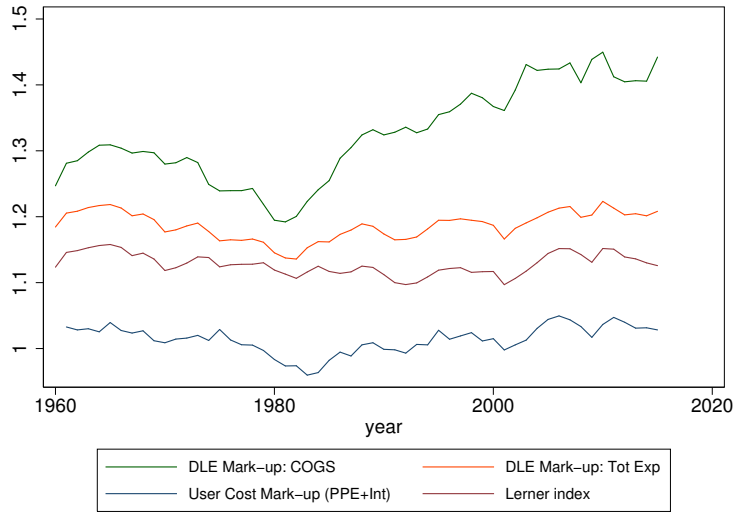
As shown, all four series exhibit an increase after 2000. However, only the DLE mark-up exhibits a substantial increase in the 1980's and 90's. The Lerner index decreases slightly from 1960 through 2000 while user cost estimates increase slightly. DLE mark-ups based on total expenses behave similar to the Lerner Index (as expected, given the use of a constant elasticity). Indeed, excluding measurement-error correction in the DLE approach and depreciation in the calculation of the Lerner index, the Lerner index is essentially equivalent to the DLE mark-up with total expenses.

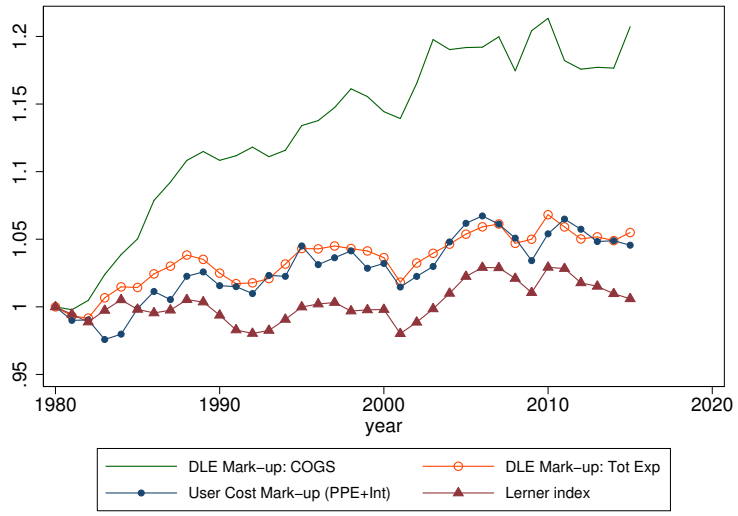
[32] They also report results using a Translog production function, which yields similar conclusions

[33] Granted, not all such expenses are variable so this is more an empirical test than a theoretically valid implementation of the approach

[34] The differences in samples explain some of the differences compared to DLE results

Figure 13: Four measures of mark-ups





Notes: Annual data primarily from Compustat.

B Data Appendix

We use a wide range of data throughout the paper as summarized in Table 1. This appendix describes the data sources, mappings and definitions. We provide a detailed description of the U.S. datasets in Section B.1; and an overview of the Europe dataset in Section B.2. For additional details on the Europe dataset, please refer to Dottling et al. (2017).

B.1 U.S. data

This section describes the main data sources and definitions used for U.S. analyses. It follows a similar structure as Table 1. Sections B.1.1 and B.1.2 describe aggregate and industry-level data. Section B.1.3 describes firm-level data. Section B.1.4 describes several measures of concentration along with a variety of analyses used to ensure consistency across them. Section B.1.5 describes import and NTR gap data used for the China analyses. Last, Section B.1.6 describes datasets containing industry-level measures of productivity and production structure (namely, the NBER-CES and U.S. KLEMS datasets).

B.1.1 Aggregate data

We gather aggregate and sector-level (Non-Financial Corporate and Non-Financial Non-Corporate) data on funding costs, profitability, investment and market value for the U.S. Economy from the U.S. Flow of Funds accounts through FRED. These data are used in the analyses discussed in the introduction; and to reconcile and ensure the accuracy of more granular data.

The following data series and definitions are used:

Net investment. For the NFC sector, net investment is defined as gross fixed capital formation minus consumption of fixed capital (series NCBGFCA027N minus NCBCFCA027N). For the NFNC sector, we use gross fixed investment minus consumption of fixed capital (series NNBGFNQ027S minus NNBCCFQ027S).

Capital stock. The capital stock is defined as the sum of equipment, intellectual property, residential and non-residential structures. For the NFC sector, these are series ESABSNNCB, NCBNIPPCCB, RCVSRNWMVBSNNCB and RCSNNWMVBSNNCB. For NFNC sector, ESABSNNB, NNBNIPPCCB, RCVSRNWBSNNB and RCVSNWBSNNB.

Operating Surplus. Net operating surplus is sourced directly from series NCBOSNQ027S and NNBBOSA027N for the NFC and NFNC sectors, respectively

Tobin's Q. Tobin's Q for the non-financial corporate sector is defined as

$$Q = \frac{V^e + (L - FA) - Inventories}{P_k K} \tag{6}$$

where V^e is the market value of equity (NCBCEL), L are the liabilities (TLBSNNCB); FA are financial assets (TFAABSNNCB); and $P_k K$ is the replacement cost of capital (sum of the four NFC capital series listed above). Inventories are based on series IABSNNCB.

B.1.2 Industry data

Dataset. Industry-level investment and profitability data are gathered from the Bureau of Economic Analysis (BEA). Investment data is based on Section 3 tables, which include current-cost and chained values for the net stock of capital, depreciation and investment. Note that BEA I and K include intangible assets (i.e., software, R&D, and some intellectual property), not just tangible capital. Value added and output data – including gross output, gross operating surplus, compensation and taxes – are from the GDP by Industry accounts (files GDPbyInd_GO_1947-2016 and GDPbyInd_VA_1947-2015).

Definitions. We define industry-level gross investment rates as the ratio of 'Investment in Private Fixed Assets' to lagged 'Net Stock of Private Fixed Assets'; depreciation rates as the ratio of 'Depreciation of Private Fixed Assets' to lagged 'Net Stock of Private Fixed Assets'; and net investment rates as the gross investment rate minus the depreciation rate. Both Current-Cost and Chained-Value investment/depreciation amounts are available. We use current-cost amounts to compute Net Investment Rates, and chained quantity indices for industry-level regressions of $log(K)$.[35] Investment rates are computed across all asset types, as well as separating intellectual property from structures and equipment.

Segments. Industry-level investment data is available for 63 granular industry groupings from the BEA. These are grouped into 47 categories to ensure all groupings have material investment; reasonable Compustat coverage; and yield stable investment and concentration time series. In particular, we group industries to ensure each group has at least ∼10 firms, on average, from 1990 - 2015 and it contributes a material share of investment. We exclude Financials and Real Estate; and also exclude Utilities given the influence of government actions in their investment. Last, Management is excluded from most regressions because there are no companies in Compustat that map to this category. This leaves 43 industry groupings for most of our analyses. All other datasets are mapped into these 43 industry groupings. The groupings are summarized in Table 8.

[35] Our results are generally robust to using chained quantity indices instead of current costs when computing net investment rates

Table 8: Mapping of BEA industries to segments

BEA code	Industry	Mapped segment	Capital (2014)	% of total
721	Accommodation	Acc_accommodation	358.9	2.2%
722	Food services and drinking places	Acc_food	249.2	1.5%
561	Administrative and support services	Adm_and_waste_mgmt	189.2	1.2%
562	Waste management and remediation services	Adm_and_waste_mgmt	102.3	0.6%
110	Farms	Agriculture	567.7	3.5%
113	Forestry, fishing, and related activities	Agriculture	62.3	0.4%
713	Amusements, gambling, and recreation industries	Arts	163.7	1.0%
711	Performing arts, spectator sports...	Arts	159.9	1.0%
230	Construction	Construction	284.6	1.7%
334	Computer and electronic products	Dur_Computer	506.3	3.1%
335	Electrical equipment, appliances...	Dur_Electrical	73.5	0.5%
333	Machinery	Dur_Machinery	234.4	1.4%
337	Furniture and related products	Dur_Furniture	22.8	0.1%
338	Miscellaneous manufacturing	Dur_Misc	115.1	0.7%
336	Motor vehicles, bodies and trailers, and parts	Dur_Transportation	383.7	2.4%
321	Wood products	Dur_Wood	42.6	0.3%
327	Nonmetallic mineral products	Dur_nonmetal	87.1	0.5%
331	Primary metals	Dur_prim_metal	165.5	1.0%
332	Fabricated metal products	Dur_fab_metal	175.3	1.1%
610	Educational services	Educational	557.7	3.4%
521	Federal Reserve banks	Finance	Omitted	
522	Credit intermediation and related activities	Finance	Omitted	
523	Securities, commodity contracts, and investments	Finance	Omitted	
524	Insurance carriers and related activities	Finance	Omitted	
525	Funds, trusts, and other financial vehicles	Finance	Omitted	
622	Hospitals	Health_hospitals	916.1	5.6%
623	Nursing and residential care facilities	Health_hospitals	94.6	0.6%

B.1.3 Firm-level data

Dataset. Firm-level data is primarily sourced from Compustat, which includes all public firms in the U.S. Data is available from 1950 through 2016, but coverage is fairly thin until the 1970's. We exclude firm-year observations with assets under $1 million; with negative book or market value; or with missing year, assets, Q, or book liabilities.[36] In order to more closely mirror the aggregate and industry figures, we exclude utilities (SIC codes 4900 through 4999), real estate and financial firms (SIC codes 6000 through 6999); and focus on U.S. incorporated firms.

It is worth noting that firm- and industry-data are not readily comparable. They differ in their definitions of investment and capital, and in their coverage. As a result, we spent a fair amount

[36] These exclusion rules are applied for all measures except firm age, which starts on the first year in which the firm appears in Compustat irrespective of data coverage

Table 8: Mapping of BEA industries to segments (cont'd)

BEA code	Industry	Mapped segment	Capital (2014)	% of total
621	Ambulatory health care services	Health_other	352	2.2%
624	Social assistance	Health_other	65.4	0.4%
514	Information and data processing services	Inf_data	168.3	1.0%
512	Motion picture and sound recording industries	Inf_motion	287.8	1.8%
511	Publishing industries (includes software)	Inf_publish	196.5	1.2%
513	Broadcasting and telecommunications	Inf_telecom	1352.5	8.3%
550	Management of companies and enterprises	Mgmt	401.4	2.5%
212	Mining, except oil and gas	Min_exOil	186.5	1.1%
211	Oil and gas extraction	Min_Oil_and_gas	1475.2	9.1%
213	Support activities for mining	Min_support	142	0.9%
325	Chemical products	Nondur_chemical	900.1	5.5%
311	Food and beverage and tobacco products	Nondur_food	336.4	2.1%
313	Textile mills and textile product mills	Nondur_textile	40.4	0.2%
315	Apparel and leather and allied products	Nondur_apparel	17.5	0.1%
322	Paper products	Nondur_paper	120.7	0.7%
323	Printing and related support activities	Nondur_printing	49.4	0.3%
326	Plastics and rubber products	Nondur_plastic	104.2	0.6%
324	Petroleum and coal products	Nondur_petroleum	221	1.4%
810	Other services, except government	Other_ex_gov	619.5	3.8%
541	Legal services	Prof_serv	42.6	0.3%
541	Computer systems design and related services	Prof_serv	74.3	0.5%
541	Miscellaneous professional, scientific, and technical services	Prof_serv	477.6	2.9%
531	Real estate	Real Estate	Omitted	
532	Rental and leasing services and lessors of intangible assets	Real Estate	Omitted	
44R	Retail trade	Retail_trade	1236.4	7.6%
481	Air transportation	Transp_air	249.1	1.5%
484	Truck transportation	Transp_ground	143.6	0.9%
485	Transit and ground passenger transportation	Transp_other	44.8	0.3%
487	Other transportation and support activities	Transp_other	132.6	0.8%
493	Warehousing and storage	Transp_other	46	0.3%
486	Pipeline transportation	Transp_pipeline	227.3	1.4%
482	Railroad transportation	Transp_rail	405.7	2.5%
483	Water transportation	Transp_other	45.6	0.3%
220	Utilities	Utilities	Omitted	
420	Wholesale trade	Wholesale_trade	590.1	3.6%

of time simply reconciling the various data sources. We refer the reader to **?** for details on the reconciliation and validation exercises.

We supplement Compustat with two sources (available through WRDS):

1. I/B/E/S: We gather EPS analyst forecasts from I/B/E/S. Forecasts are used to (i) control for projected growth in our noisy entry estimate and (ii) estimate the cost of capital in the mark-up estimates reported in the Appendix A.[37]

2. Intangible Capital Estimates: Last, we gather firm-level intangible capital estimates as defined in Peters and Taylor (2016)

Firms are mapped to BEA industry segments using the NAICS Level 3 mapping outlined by the BEA. When NAICS codes are not available, firms are mapped to the most common NAICS category among those firms that share the same SIC code and have NAICS codes available. Firms with an 'other' SIC code (SIC codes 9000 to 9999) are excluded from industry-level analyses because they cannot be mapped to an industry.

Firm-level data is used for two purposes:

- First, we aggregate firm-level data into industry-level metrics and use the aggregated quantities to explain industry-level investment behavior . We consider the aggregate (i.e., weighted average), the mean and the median for all quantities, and use the specification that exhibits the highest statistical significance. We require at least 5 firms in a given industry-year pair to include a given observation in industry-level analyses (all firms are included in firm-level analyses, irrespective of the number of firms in a given industry-year).

- Second, we use firm-level data to analyze the determinants of firm-level investment through panel regressions. We compute a wide range of financial measures, including investment, cash flow, operating surplus, etc. The main variables are discussed in the following section; with additional details on the sample selection, variable definitions and data quality tests available in **?**.

Firm-level Definitions.

Capital. We consider three measures of capital: net Property, Plant and Equipment (item PPENT), Intangible capital at replacement cost (item K_INT from Peters and Taylor (2016)) and total capital (PPENT + K_INT). For regressions of K, we deflate capital measures using the Fixed investment: Nonresidential (implicit price deflator) (series A008RD3A086NBEA)

[37] IBES provides the consensus of all available forecasts as of the middle (the Thursday following the second Friday) of each month. IBES data is mapped to Compustat GVKEY through a two-step approach. First, we use the header map between GVKEY and IBES Ticker provided in the Compustat Security table (IBTIC variable). Then, for those GVKEYs that have missing IBTIC in Compustat and have a valid PERMNO, the link is supplemented with additional historical GVKEY-IBES ticker links as follows: we first merge the rest of GVKEYS with PERMNOs on a historical basis using CRSP-Compustat Merged Database. Then, we bring in additional IBES Tickers from the IBES-PERMNO link (we use the WRDS ICLINK and CIBESLINK applications).

Employment. We use Compustat field EMP as a measure of employment

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.

Entry and Exit. Entry is defined as the first year in which a firm (GVKEY) appears in Compustat and does not violate our exclusion restrictions. Enforcing exclusion restrictions has a minimal effect on our results, and ensures that our entry/exit rates map to all other analyses. Exit is defined as the last year available for a given GVKEY. We differentiate across exit types using field DLRSN, which is equal to 1 when a firm exits due to M&A.

Q. Firm-level Q is defined as the ratio of market value to total assets (AT). We compute market value as the market value of equity (ME) plus total liabilities (LT) and preferred stock (PSTK), where the market value of equity (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). The resulting aggregate and mean Q from Compustat closely mirror the Flow of Funds Q (see **?**).

Lerner Index. We follow Grullon et al. (2016) and define the Lerner Index as operating income before depreciation minus depreciation divided by sales.

B.1.4 Measures of Concentration.

This section describes our approach to measuring concentration across U.S. industries. We start by providing additional details on our calculation of import-adjusted Herfindahls (relative to the main body). We then discuss a variety of analyses and results used to gain comfort around the Herfindahl estimates. Last, we discuss additional measures of concentration (including Compustat and Census Concentration Ratios) used to complement Herfindahls.

Herfindahls. Our main measure of competition is the import-adjusted Sales Herfindahl $\left(HHI_t^k\right)$ at the BEA segment-level k. This measure is constructed differently across sectors depending on data availability and the importance of foreign competition.

Manufacturing. For manufacturing industries, HHI_t^k is estimated based on the data of Feenstra and Weinstein (2017). In particular, the replication files available at the author's website include Herfindahls at the country- and 4-digit Harmonized System (HS-4) code, from 1992 to 2005. For the U.S., these Herfindahls are based on the U.S. Economic Census. For the remaining countries, they are based on import data at the firm- and HS-4 sector-level.

We start by aggregating across countries c to calculate the *Overall Herfindahl Index* for HS-4 sector j

$$HHI_t^j = \sum_{c \in C_t} \sum_{i \in I_{ct}} (s_{cit}^j)^2 = \sum_{c \in C_t} \sum_{i \in I_{ct}} (s_{it}^{cj})^2 (s_{ct}^j)^2 = \sum_{c \in C_t} HHI_t^{cj} (s_{ct}^j)^2$$

where i denotes firms, s^j_{cit} denotes firm i's share in total US-sales for sector j; s^j_{ct} denotes the total import share from country c in sector j; and $s^{cj}_{it} = s^j_{cit}/s^j_{ct}$ denotes firm i's share within the exports of country c in sector j.

Next, we use the correspondence of Pierce and Schott (2012) to map HS-4 sectors to NAICS-6 sectors,[38] and then map NAICS-6 sectors to BEA segments (which roughly correspond to NAICS-3 segments). NAICS-6 is generally more aggregated than HS-4 so that, most of the time, HS-4 codes map one-to-one to BEA segments. But this is not always the case.[39] Given that more granular data is not available, we assume Herfindahls are constant for all HS-4 sectors j within BEA segment k and compute HHI^{jk}_t – the Herfindahl across firms in HS-4 sector j within BEA segment k:

$$HHI^{jk}_t = HHI^j_t / \sum_{j \in I_k} (s^k_{jt})^2.$$

where $j \in I_k$ denotes HS-4 sectors within BEA segment k and s^k_{jt} denotes the share of HS-4 sector j within BEA segment k. When HS-4 sector j maps only to BEA segment k , $HHI^{jk}_t = HHI^j_t$.

We then aggregate HHI^{jk}_t to the BEA segment-level k using

$$\sum_{j \in I_k} HHI^{jk}_t (s^k_{jt})^2 = \sum_{j \in I_k} \sum_{i \in K_j} (s^{jk}_{it})^2 (s^k_{jt})^2 = \sum_{j \in J_k} (s^k_{ijt})^2 = HHI^k_t$$

where s^{jk}_{it} is the share of firm i in sector j within BEA segment k; and $s^k_{ijt} = s^{jk}_{it} s^k_{jt}$ is the share of firm i within HS-4 segment j that belongs to BEA segment k.[40]

Unfortunately, granular import data from Feenstra and Weinstein (2017) is only available from 1992 to 2005 – and we would like to analyze investment over a broader period. In addition, FW Herfindahls are based on SIC segments before 1997 and NAICS segments afterwards, which results in a jump in HHI^k_t for some series. We control for the jump by subtracting the 1997 change in HHI^k_t from all HHI^k_t series after 1998.[41] We then extend the time series through a regression of the form:

$$log\left(HHI^k_t\right) = log\left(HHI^{k,CPraw}_t\right) + log\left(s^{CP}_{kt}\right) + \alpha^k + \varepsilon_{kt} \tag{7}$$

where $HHI^{k,CPraw}_t$ denotes the raw Herfindahl from Compustat firms in our sample and s^{CP}_{kt} denotes the share of sales of Compustat firms as a percent of total US output plus imports:

$$s^{CP}_{kt} = \frac{\sum_{i \in k} sale_{it}}{gross\, output_k + imports_k}$$

[38] In particular, we use the implied mapping in the import files used for the China analyses

[39] For instance, HS codes 9615903000 and 9615904000 map to NAICS 339 and 326, respectively, under the Pierce and Schott (2012) correspondence.

[40] The derivations in this section closely follow Feenstra and Weinstein (2017) – particularly equations A4 and A5.

[41] We also find a jump in the Durable - NonMetal sector between 1996 and 1997, driven by the introduction of HS-4 code 3824 which carries large domestic sales and a high Herfindahl. This appears unintuitive so we subtract the associated change from this industry's series. Our results are robust to keeping the raw data.

where $\sum_{i \in k} sale_{i,t}$ denotes the sum of sales across all Compustat firms i that belong to BEA segment k in our sample; $gross\,output_k$ denotes the total output (i.e., sales) of BEA segment k (as measured by the GDP by Industry tables) and $Imports_k$ denotes total imports in BEA segment k (sourced from Peter Schott's website).We use gross output plus imports in the denominator (rather than absorption) because Compustat sales are inclusive of exports. Still, $s^{CP}_{j,t}$ exceeds 1 for a few industries. This is likely due to two issues in mapping Compustat to the Census: (i) Compustat includes total sales irrespective of the location of production, while the National Accounts include only production in the U.S.; and (ii) National Accounts map output to industries within firms (i.e., the sales of a firm may be mapped to more than one industry), while in Compustat entire firms are mapped to industries. To mitigate these issues, we cap s^{CP}_{kt} at the minimum of 0.9 and the U.S. share of output in a given industry:

$$s^{BEA}_{kt} = \frac{gross\,output_k}{gross\,output_k + imports_k}$$

We also confirm that using $s^{BEA}_{kt} = {}^{output_k}/_{(output_k + imports_k)}$ instead of s^{CP}_{kt} across all estimations yields consistent results.

The resulting Herfindahl estimate $\left(HHI^k_t\right)$ following FW tends to be lower than the Compustat Herfindahl, even after adjusting for the share of Compustat sales. This is likely because of a lower concentration of foreign/non-Compustat firms compared to Compustat. Most of our regressions include fixed effects, so this is not an issue. However, for columns 1-2 in Table 6 as well as some Figures, the level of the HHI^k_t matters. We therefore add a constant across all manufacturing segments, to match the average level of HHI^k_t to that of $HHI^{k,CPadj}_t$ across all manufacturing industries.

Mining and Agriculture. For mining and agriculture, we would ideally follow the approach of Feenstra and Weinstein (2017). However, neither US census concentration measures nor firm-level import data are available. Instead, we use the Compustat-based US Herfindahl HHI^{CP}_{kt} and estimate HHI^k_t as follows: consider an industry with x firms in Compustat and N firms globally, all with equal shares of the U.S. market. The Compustat share of output is $s^{CP} = \frac{x}{N}$, and the Compustat-based Herfindahl $HHI^{CP} = \frac{1}{x}$. Under these assumptions, the import-adjusted Herfindahl could be computed as

$$\begin{aligned} HHI &= \frac{1}{N} \\ &= HHI^{CP} \times s^{CP} \end{aligned}$$

which is true as long as the Non-Compustat Herfindahl scales down with sales.[42] This is a strong assumption – but it is hard to justify a different one. Thus, we set $HHI^k_t = HHI^{CPraw}_{kt} \times s^{CP}_{kt}$

[42] Formally, we have $HHI = HHI^{CP}(s^{CP})^2 + HHI^{NonCP}(s^{NonCP})^2 = HHI^{CP} \times s^{CP}$ if and only if $HHI^{NonCP} = HHI^{CP}\frac{s^{cp} - \left(s^{CP}\right)^2}{\left(s^{NonCP}\right)^2}$.

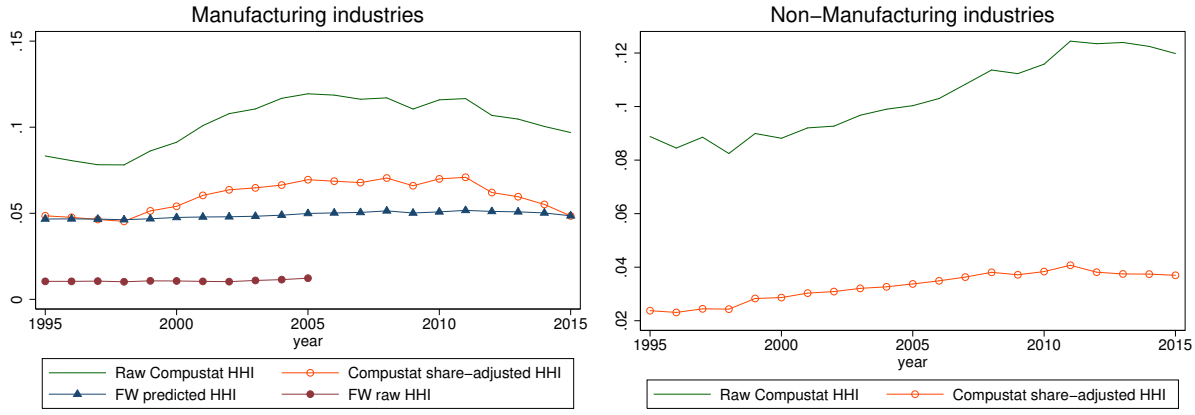
where s_{kt}^{CP} is the share of Compustat sales in US output plus imports as defined above. We refer to this measure of Herfindahls as the "Compustat-adjusted" Herfindahl (HHI_{kt}^{CPadj}).

Services. For service sectors, data on imports is not available but it also constitutes a much smaller share of sales. We follow a similar approach as for mining and agriculture, except that we set imports to zero. Namely, we estimate the industry Herfindahl as $HHI_t^k = HHI_{kt}^{CP} \times s_{kt}^{CP}$, where the share of sales includes only domestic output

$$s_{kt}^{CP} = \frac{\sum_{i \in k} Sale_{it}}{gross\, output_k}$$

Figure 14 replicates Figure 3 from the main body but separating manufacturing (left) and non-manufacturing (right) industries. We include the "raw" FW-Herfindahl on the manufacturing plot, as well as the predicted FW HHI based on regression (7) and following the addition of a constant. The Raw Compustat HHI for manufacturing industries appears to have increased, but this is entirely due to foreign competition. The import adjusted series is essentially flat. By contrast, both the Raw and adjusted series for non-manufacturing industries have increased.

Figure 14: Weighted Average Herfindahls by Sector



Notes: Annual data. Figure plots the weighted average of alternate HHI measures, by sector. The Raw Compustat HHI is the sum of squared Compustat market shares. The Compustat share-adjusted HHI adjusts for the Compustat share of domestic output plus imports. The FW predicted HHI is computed as described in the text. Last, the FW raw HHI is the FW HHI after subtracting the change from 1997 to 1998.

Robustness Tests. Our estimation of import-adjusted Herfindahls relies on two main assumptions: (i) the use of Herfindahls following Feenstra and Weinstein (2017) for manufacturing industries and (ii) the use the Compustat share of sales to adjust Compustat Herfindahls into "U.S" Herfindahls for non-manufacturing industries. This section presents a series of analyses used to test these assumptions.

To begin with, Table 9 shows the pair-wise correlations (in levels and 5Y changes) between five alternate measures of Herfindahls:

Table 9: Pair-wise Correlations between CP and FW Herfindahls

Table shows the pair-wise correlation matrix between several Herfindahl measures as described in the text. Cells including FW Herfindahls include only manufacturing industries. * = significant at 1% level.

		CP-Raw	CP share	US share	FW	FW-pred
Levels	CP-Raw	1				
	CP- Adj by CP share	0.6856*	1			
	CP Adj by US share	0.9871*	0.6963*	1		
	FW	0.5323*	**0.8051***	0.5759*	1	
	FW-pred	0.4313*	0.8001*	0.5045*	**0.9785***	1
5Y changes	CP-Raw	1				
	CP- Adj by CP share	0.6871*	1			
	CP Adj by US share	0.9889*	0.6906*	1		
	FW	0.3717*	**0.6587***	0.3933*	1	
	FW-pred	0.6867*	0.9023*	0.7212*	**0.6642***	1

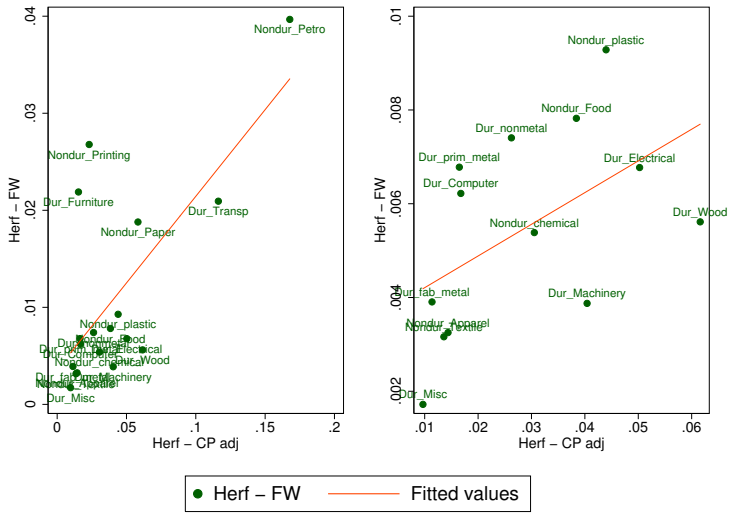
1. **CP-Raw** is the raw Compustat Herfindahl, estimated as the sum of squared market shares across all Compustat firms in a given industry.
2. **CP- Adj by CP share** adjusts CP-Raw based on the Compustat share of sales (s^{CP})
3. **CP Adj by US share** adjusts CP-Raw based on the US share of sales (s^{BEA})
4. **Raw FW** is the Feenstra-Weinstein Herfindahl $\left(HHI_t^k\right)$, where we subtract the change in 1997 to account for the shift from SIC to NAICS segments
5. **Predicted FW** is the predicted Feenstra-Weinstein Herfindahl based on regression 7

As shown, all measures of Herfindahls are strongly correlated. Two additional items are worth highlighting. First, the CP-adjusted Herfindahl exhibits the strongest correlation between FW-Herfindahls and Compustat Herfindahls – both in levels and changes. This is expected, as the adjustment better accounts for the limitations of using Compustat. It also supports our choice of using CP-adjusted Herfindahls for non-manufacturing industries. Second, the FW-predicted Herfindahl is highly correlated with the actual FW-Herfindahl, suggesting that our extrapolation is robust.

Focusing on manufacturing industries, Figure 15 plots the FW Herfindahl against the CP-adjusted Herfindahl as of 1993, to provide a more detailed comparison between them. The left plot includes all industries, while the right plot includes only industries with a FW-Herfindahl below 0.01 (given the clustering). FW and CP-adjusted Herfindahls exhibit positive (and significant) relationship. As noted above, the CP-adjusted Herfindahls are substantially higher – in level – than the FW Herfindahls, which justifies the addition of a constant.

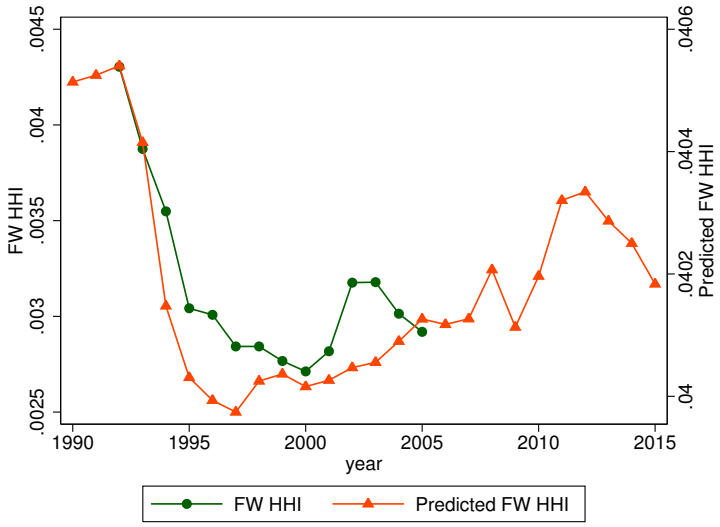
Figure 16 compares the "raw" FW Herfindahl to the predicted FW Herfindahl based on regression 7 for a sample industry (Durable Machinery). The two series behave similarly. This is true for the majority of industries.

Figure 15: Comparison of FW and CP share-adjusted Herfindahls



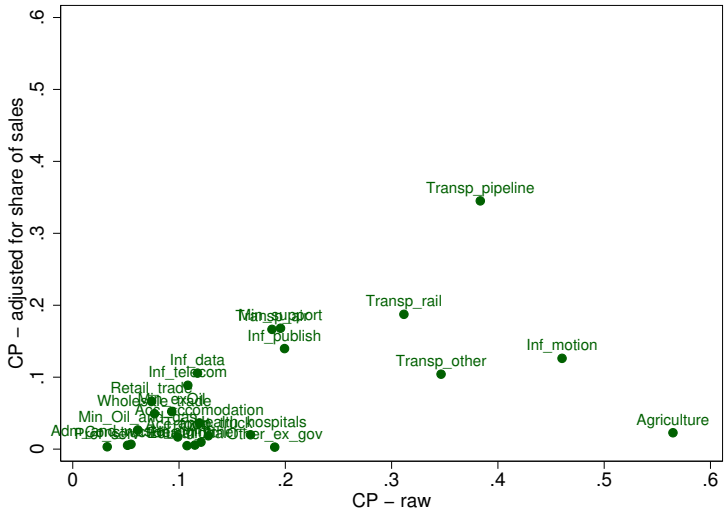
Notes: Annual data as of 1993. Figure plots the FW Herfindahl against the Compustat Herfindahl adjusted for the share of Compustat sales. The left plot includes all industries, while the right plot includes only industries with a FW-Herfindahl below 0.01. As shown, both series are strongly correlated.

Figure 16: Raw and predicted FW Herfindahl: Dur Machinery industry



Notes: Annual data. Figure plots the raw and predicted FW Herfindahl for the Durable Manufacturing - Machinery industry, computed as described in the text.

Figure 17: Comparison of Raw and Adjusted CP Herfindahls



Notes: Annual data as of 2015. Figure shows the "raw" Compustat Herfindahl against the Herfindahl adjusted for the share of Compustat sales. The share adjustment has a material implications, particularly for industries with low coverage in Compustat that otherwise exhibit very large Herfindahls. See text for additional details.

Moving on to non-manufacturing industries, Figure 17 plots the CP-adjusted Herfindahl against the Raw Compustat Herfindahl. As shown, adjusting for the Compustat share of sales has material implications – particularly for industries where Compustat provides limited coverage (e.g., Agriculture). In those cases, the Compustat Herfindahl is too high by construction (because there are few firms), so that the share adjustment is critical.

Other Measures of Competition. Our primary measure of competition is the import adjusted Herfindahl, but we also gather and report selected results with Census CRs (sourced from https://www.census.gov/econ/concentration.html). These data are available every five years, at the NAICS Level 3 to Level 6 granularity for a subset of industries. They include the share of sales held by the top 4, 8, 20 and 50 firms in each industry; and are based on SIC categories up to 1992, and NAICS categories from 1997 to 2012. We use NAICS Level 6 CRs for the TFP analyses of manufacturing industries that rely on the NBER-CES database. We use NAICS Level 3 CRs (mapped to BEA segments by taking the weighted average by sales) in the TFP analyses based on US KLEMS and in selected regression analyses reported in Appendix C. We also gather SIC-based concentration ratios from 1982 to 1992 and use them to create aggregate trends such as those reported in Figure 1. In particular, we were able to find SIC-4 CRs for manufacturing industries as of 1992; and for non-manufacturing industries as of 1987 and 1992 from the Economic Census. We obtained the weighted average manufacturing concentration ratio as of 1982 from Pryor (2001). We do not use SIC-based CRs in regression analyses because all our analyses are based on NAICS segments and are primarily focused on recent years.

Census CRs can also help us validate the accuracy of Compustat Concentration measures. In par-

Table 10: Pair-wise Correlations between Compustat and Census Concentration Measures

Table shows pair-wise correlations between Compustat (rows) and Census (columns) measures of competition across BEA segments. * = significant at 5% confidence.

	Levels		5-year Changes	
	$CR4^{Cen}$	$CR8^{Cen}$	$CR4^{Cen}$	$CR8^{Cen}$
HHI^{CP}	0.2527*	0.2540*	0.2596	0.2241
$CR4^{CP}$	0.3449*	0.3597*	0.3465*	0.3498*
$CR4^{CP}$ *Adj*	**0.8146***	**0.8655***	**0.5859***	**0.5618***
$CR8^{CP}$	0.3025*	0.3226*	0.2374	0.2415
$CR8^{CP}$ *Adj*	**0.7999***	**0.8532***	**0.4556***	**0.3956***

ticular, Table 10 shows the pairwise correlations between Compustat (rows) and Census (columns) measures of concentration. For Compustat, we include the "raw" Compustat Herfindahl, as well as two measures of the top 4 and top 8 firm CRs. CRx^{CP} is the raw top 'X' firm CR while CRx^{CP} *Adj* adjusts for the share of Compustat sales (e.g., $CR4^{CP}$ $Adj = CR4^{CP} \times s^{CP}$). For the Census, we include the Top 4- and Top 8-firm CR. As shown, Compustat and Census competition measures are strongly correlated in both levels and changes – particularly once adjusting for the Compustat share of sales.

Conclusion on Measures of Concentrations. The analyses described in this section suggest that our measures of concentration are – to a reasonable extent – robust and internally consistent. Still, none of these measures is perfect. We perform extensive sensitivity analyses to adjustments in the calculation of import-adjusted Herfindahls (e.g., using s_{kt}^{BEA} instead of s_{kt}^{CP}). We also note that many of our results are not entirely dependent on the choice of Herfindahl.

Out of our evidence of Decreasing Domestic Competition – the comparison of U.S. and European Industries relies on domestic Herfindahls, which are easier to measure. And we also validate that our results are robust to two alternate measures of competition in each region: Compustat and Census CRs in the U.S.; and BvD and CompNET in Europe. Our results related to rising concentration and TFP are again based on domestic Census Concentration measures, which are the correct unit of comparison for domestic TFP growth. Our finding that leaders decrease investment and R&D in concentrating industries is robust to using import-adjusted Herfindahls as well as census-based concentration measures.

Similarly, our evidence of decreased investment is not entirely dependent on Herfindahls. Increased competition in the China Natural Experiment is measured by NTR gaps, not Herfindahls or CRs. Noisy Entry is a significant predictor of investment individually – in addition to a good instrument for Herfindahls. And Herfindahls are only used as controls in our M&A boom analyses.

B.1.5 China import-competition data

Data on international trade is sourced from the UN Comtrade Database[43] and Peter Schott's website.

UN Comtrade data is used by Autor et al. (2016) (among others) and includes bilateral imports by six-digit Harmonized Commodity Description and Coding System (HS) products. Data for a consistent set of countries is available from 1991 to 2014. We map these data to six-digit NAICS codes by applying the crosswalk in Pierce and Schott (2012), which maps 10-digit HS products to six-digit NAICS industries.

Following Autor et al. (2016), we define the import penetration ratio for industry j as:

$$\Delta IP_{jt}^{US} = \frac{\Delta M_{jt}^{US}}{Y_{j,91} + M_{j,91} - E_{j,91}} \tag{8}$$

where $\Delta M_{j\tau}^{US}$ denotes the change in imports from China from 1991 to t; and $Y_{j,91} + M_{j,91} - E_{j,91}$ denotes the initial absorption (defined as output, $Y_{j,91}$, plus imports, $M_{j,91}$, minus exports, $E_{j,91}$). $Y_{j,91}$ is sourced from the NBER-CES database; while $M_{j,91}$ and $E_{j,91}$ are sourced from the UN Comtrade Database (and measure U.S. imports and exports with the rest of the world). Only NAICS level 6 industries where data is available across all sources are included in the analyses.[44] We use 1991 as our benchmark year because data is available across a broad sample of countries starting that year.

We also compute import penetration from China to eight other high-income countries:

$$\Delta IP_{jt}^{OC} = \frac{\Delta M_{j\tau}^{OC}}{Y_{j,91} + M_{j,91} - E_{j,91}}$$

where $\Delta M_{j\tau}^{OC}$ denotes the change in imports from China in industry j during year t to eight other high-income countries; while the denominator is the same as above.[45]

Last, we obtain Non-Normal-Trade-Relations tariff gaps and Census-based import data from Peter Schott's website. NTR tariff gaps are used in Pierce and Schott (2016) and are defined for NAICS level 6 industries.[46]

[43] http://comtrade.un.org/db/default.aspx

[44] The main concern with this is that some industry segments in the NBER-CES have no representation in Compustat and/or no import data from UN Comtrade. To mitigate this, we repeat all tests using an alternate measure of import penetration that does not rely on the NBER-CES database: the ratio of changes in Chinese imports to total U.S. imports as of 1991 $\left(\Delta IP_{j\tau}^{2} = \frac{\Delta M_{j\tau}^{UC}}{M_{j,91}}\right)$; and exclude all industry-level controls which are sourced from the NBER-CES database. We find consistent – and in some cases more significant – results when using this broader sample, which suggests that omitted industries due to data availability are not driving our results.

[45] Following Acemoglu et al. (2016), we use Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland as our benchmark countries. We use 1991 as the benchmark year in the denominator instead of 1988 as used by Acemoglu et al. (2016) because import and export data for the U.S. following HS codes is available only since 1991 in UN Comtrade. Prior to 1991, data is based on SITC codes; which would require an additional mapping to NAICS codes. Moreover, our focus is on the post-2000 period, so 1991 already embeds a lag to reduce endogeneity concerns.

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

B.1.6 TFP and Industry Characteristics

Last, we obtain data on industry-level productivity and production structure from the 2016 release of the NBER-CES Manufacturing Industry Database (http://www.nber.org/nberces/) and the Multifactor Productivity statistics of the U.S. Bureau of Labor Statistics (https://www.bls.gov/mfp/).

The NBER-CES database 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. We use these data in the granular TFP analyses of section 2.2 as well as to compute industry-level absorption and controls for the China analyses.

From the Multifactor Productivity statistics, we gather industry-level TFP defined over the same industries as BEA segments. We map reported TFP to our 43 BEA industry groupings by taking the value-added weighted average of TFP changes, and again use them for the TFP regressions of section 2.2.

B.2 Europe data

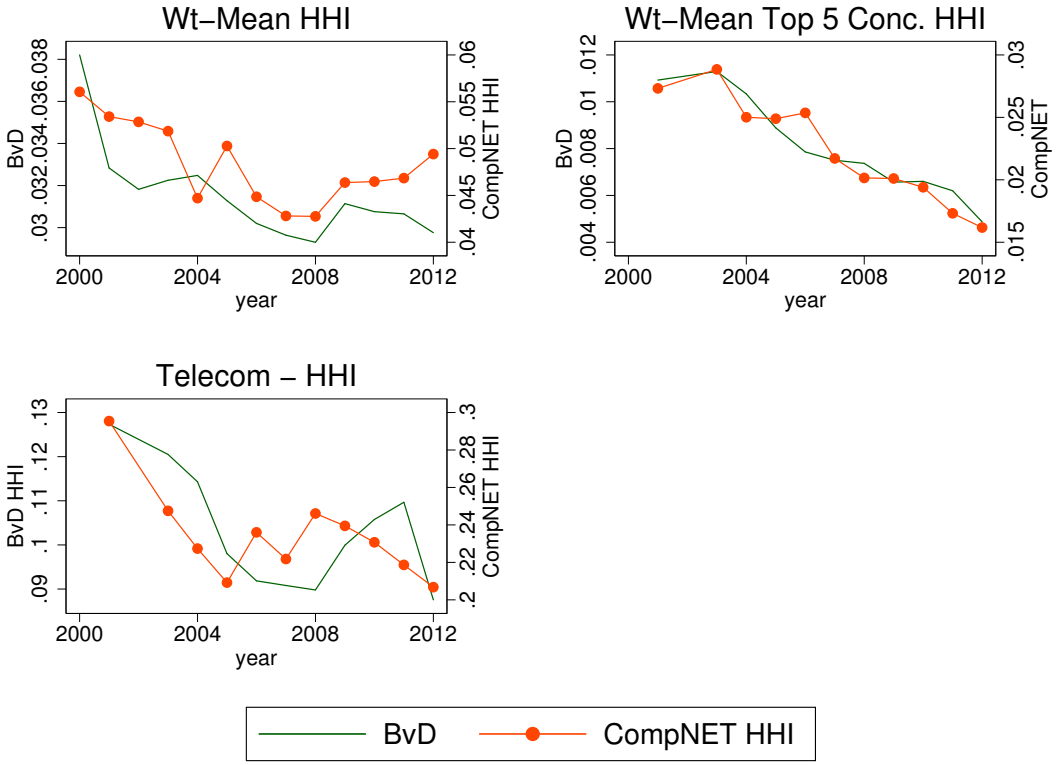
We use Europe as a benchmark in two places: while comparing the evolution of US and European industries in Section 2.1; and as a robustness test for the relationship between TFP and concentration in Section 2.2.

The analyses in Section 2.1 are based on the dataset of Dottling et al. (2017). Herfindahls are computed based on the BvD Orbis merged vintage dataset of Kalemli-Ozcan et al. (2015). Industry-level investment and capital stocks are sourced from OECD STAN; and Tobin's Q is based on firm-level data from Compustat Global and CRSP. Please refer to Dottling et al. (2017) and the associated data appendix for additional details.

BvD Orbis is used to compute Herfindahls because – compared to the US – a larger fraction of firms is held privately in European countries. So Compustat would provide an even more biased measure of concentration for Europe than the U.S. BvD Orbis contains accounting information for private as well as public firms, hence provides a robust estimate of concentration for Europe. Nonetheless, we validate the BvD Orbis concentration measures by comparing them to those of the ECB's CompNET. CompNET relies on firm-level data from a variety of sources to compute measures of concentration at the industry-year level. It aims to cover the entire non financial corporate sector across all countries.

We find that concentration measures from BvD Orbis and CompNET are strongly correlated. Herfindahls (CRs) exhibit 52% (82%) correlation at the country-industry level and 70% (83%) correlation at the EU-industry level – where EU-industry measures are computed based on the turnover weighted average across countries. Figure 18 illustrates the consistency across estimates, by plotting concentration measures based on CompNET and BvD at different levels of granularity. We include only those countries included in Dottling et al. (2017), for which data is available in CompNET (Austria, Belgium, Germany, Finland, France and Italy). The left plot shows the weighted average

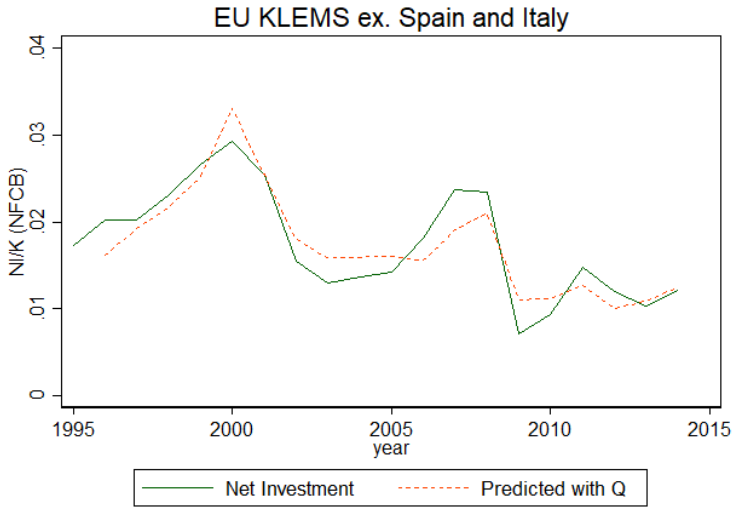
Figure 18: Comparison of BvD and CompNET Concentration Ratios



Notes: Annual data as of 2015. Figure compares BvD and CompNET Herfindahls at varying levels of granularity.

Herfindahl across all countries and industries; the middle plot shows the weighted average Herfindahl at the top 5 concentrating industries in the U.S. (the same as in Figure 6); and the right plot shows the weighted average Herfindahl of the Telecom industry. As shown, BvD and CompNET exhibit similar patterns.

Figure 19: I/K vs. Q for EU KLEMS ex. Spain and Italy



Sources: Annual data for Non-Financial Corporate sector. US data sourced from FRED. Data for European economies sourced primarily from OECD, except for Spain and Italy for which some of the data is sourced directly from the corresponding National Accounts.
Notes: Figure shows the actual and predicted net investment rate by for Non-Financial Corporate sector. Predicted series based on a regression of net investment on lagged Q from 1996 to 2009 for Europe and 1990 to 2001 for the US.

C Additional Results

This Appendix presents additional results and robustness tests. It is structured in the same way as Sections 2 and 3.

C.1 Rising Concentration Reflects Decreasing Domestic Competition

C.1.1 Comparison of US vs. Europe

This sections presents additional results related to the Comparison of U.S. and Europe. To begin with, Figure 19 plots the actual and predicted net investment rate by for Non-Financial Corporate sector for five major economies in Europe (Austria, Germany, Finland, France, Netherlands). We exclude Spain and Italy given the effect of the sovereign crisis. The predicted series is based on a regression of net investment on lagged Q from 1996 to 2009. As shown, investment in Europe appears to be in line with Q.

Next, we provide selected comparisons of the concentration, investment and mark-up patterns between the U.S. and Europe. We include the Telecom industry to align with Faccio and Zingales (2017) and the Transportation - Air industry to align with The Economist's article, "A lack of competition explains the flaws in American aviation" (April 2017). For Information - Telecom, the Herfindahl increases in the U.S. while it decreases in Europe. Similarly, investment decreases in the U.S. while it remains relatively stable in Europe. For Transportation - Air, both the Herfindahl and

Table 11: Investment by Leaders

Table shows the results of firm-level panel regressions of the log change in the stock of capital (deflated to 2009 prices) on census-based CRs. Regression from 2000 to 2015, following equation (1), except that HHI is replaced with $CR8^{Cen}$. We consider three measures of capital: PP&E, intangibles defined as in Peters and Taylor (2016) and their sum (total). Leaders measured as the two-year moving average of an indicator for a firm having market value in the top quartile of the corresponding BEA segment k. Q and log-age included as controls. As shown, leaders decrease investment with concentration, rather than increase it. Annual data, primarily sourced from Compustat. Standard errors in brackets, clustered at the firm-level. + p<0.10, * p<0.05, ** p<.01.

	(1) $\Delta log(PPE)^a$ ≥ 2000	(2) $\Delta log(Int_{PT})^b$ ≥ 2000	(3) $\Delta log(K_{PT})^{a+b}$ ≥ 2000
Q(t-1)	6.45**	3.15**	3.64**
	[0.27]	[0.14]	[0.13]
$CR8_j$(t-1)	-0.19+	-0.01	0
	[0.10]	[0.07]	[0.07]
$Leader$(t-1)	1.56	2.54	1.81
	[1.94]	[1.77]	[1.49]
$CR8_j \times Lead$	**-0.06**	**-0.18***	**-0.14***
	[0.08]	**[0.08]**	**[0.07]**
log(Age_{t-1})	-7.23**	-15.57**	-13.78**
	[1.74]	[1.01]	[0.95]
Observations	43666	43855	44004
R^2	0.05	0.08	0.09
Year FE	YES	YES	YES
Firm FE	YES	YES	YES

Price Cost Margin have increased in the U.S., while they remained stable in Europe.[47] Results for the remaining industries included in Figure 6 are similar.

C.1.2 Investment by Leaders

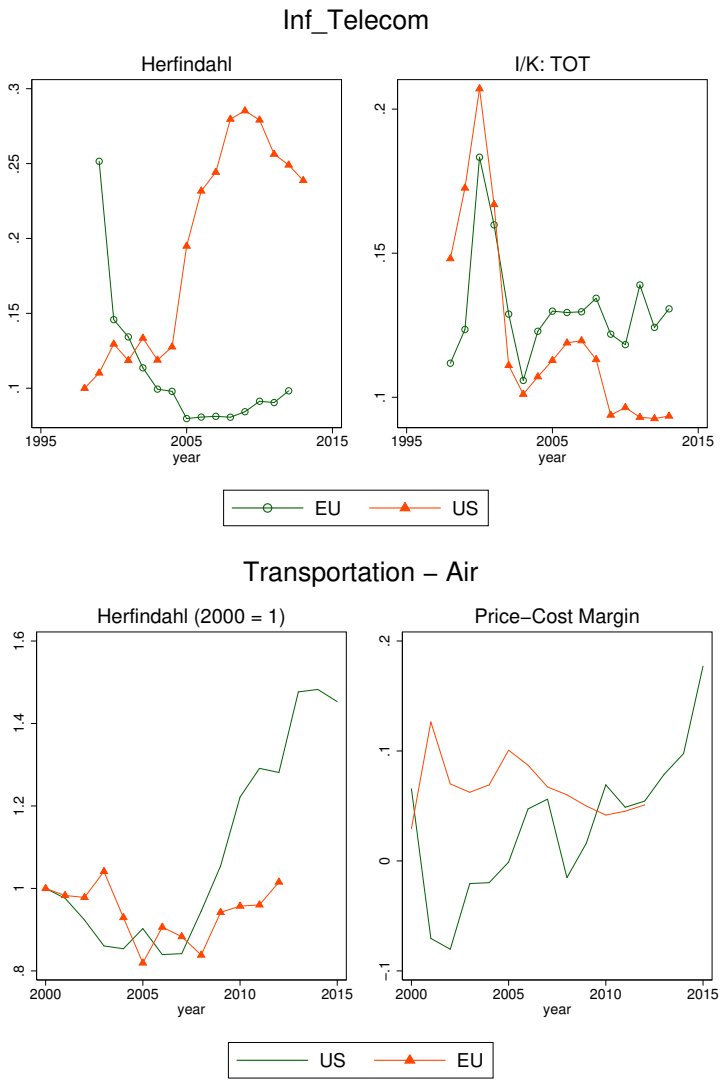
Figure 21 shows that the rise in profits is driven by leaders.[48] Laggards have not experienced an increase in profitability.

Table 11 replicates the results of Table 4 but using Census-based CRs as opposed to import-adjusted Herfindahls. In order to maintain a consistent panel regressions structure, we interpolate the 8-firm concentration ratio within each industry. As shown, results are (almost always) robust to using CRs.

[47] Note that the Herfindahl and Price Cost Margin series for Transportation - Air in Europe are based on CompNET instead of BvD Orbis/KLEMS. This is because KLEMS – and therefore Dottling et al. (2017) – combines Airlines into a broader industry containing all of Transportation and Storage. CompNET splits this industry into components and allows us to create this comparison. We normalize the Herfindahl to 1 as of 2000 for ease of comparison.

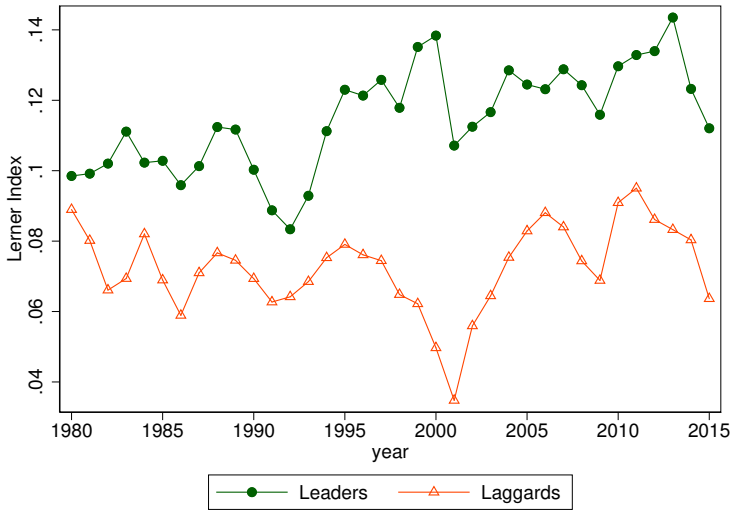
[48] For this analysis, leaders include those firms with the highest market value, which combined account for 33% of the market value in each industry and year. Laggards are those firms with the lowest market value that combined account for 33% of industry market value in each year.

Figure 20: Sample Industry-level Comparisons of U.S. and Europe



Notes: Top chart compares the evolution of the Herfindahl and investment rate in the Telecom industry for the U.S. and Europe. European series are based on the weighted average across major EU economies. Bottom charts compare the evolution of the Herfindahl and price-cost Margin in the Transportation - Air industry for the U.S. and Europe. U.S. series based on Compustat. European series based on CompNET.

Figure 21: Lerner Index of Leaders and Laggards



Notes: Annual data. Lerner Index from Compustat, defined as operating income before depreciation minus depreciation divided by sales. Bottom plot separates Leaders (firms with the highest market value that, combined, account for 33% of Market Value within each industry and year) and Laggards (firms in bottom 33% of Market Value).

C.2 DDC explains the decline in investment

C.2.1 China Import Competition

This section provides additional analysis and results related to the Chinese Natural Experiment. We discuss broad trends in firm entry, firm exit, and investment for high- and low-exposure industries that lend support to our identification strategy and results – and provide additional regression results using either the $NTRGap$ directly (as in Pierce and Schott (2016)) or import penetration instrumented by import penetration of 8 other advanced economies (as in Autor et al. (2016)).

Exploratory data analysis.

Number of firms, entry and exit rates. We begin by studying the evolution of the number of firms, firm entry and firm exit, while separating industries with 'high' (above-median) and 'low' (below-median) changes in Chinese import penetration from 1991 to 2011, $\Delta IP^{US}_{j,11}$. In particular, Figure 22 shows the change in total number of firms in industries with high and low Chinese import penetration. We normalize the number of firms to 1991. As shown, both sectors exhibit roughly the same patterns before the rise of China: the number of firms was largely flat in the 1980's, increased rapidly in the 1990's and decreased with the dot-com bubble. The patterns diverge, however, starting in the mid 2000's. The number of firms in industries with high import penetration decreased much faster than the number of firms in industries with low import penetration. Today, there are half as many firms as there were in 1995 in high-exposure industries, against nearly 80% as many in low-exposure industries

Figure 22: Number of firms by Chinese exposure (1995 = 1)



Notes: Annual data. Firm data from Compustat; import data from UN Comtrade. Manufacturing industries only, split into 'high' (above-median) and 'low' (below-median) exposure based on import penetration from 1991 to 2011.

To test the statistical significance of changes in the number of firms, we perform the following regression

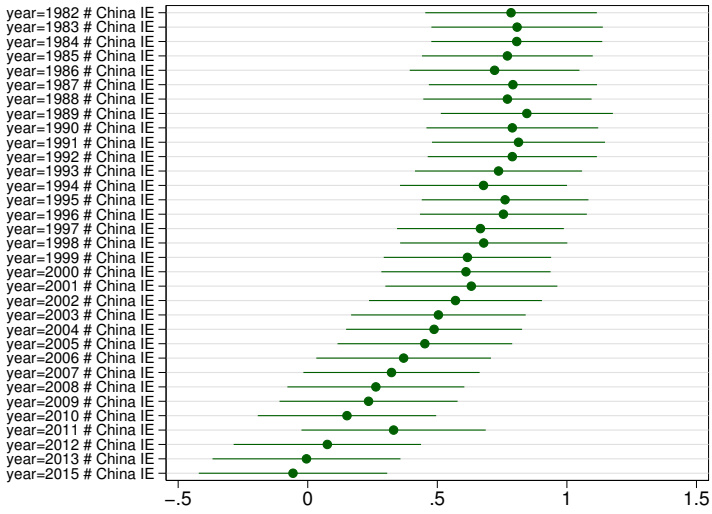
$$log(N_{j,t}) = \mu_j + \eta_t + \beta_t \Delta IP_{j,99-11} \times 1\{year\} + \varepsilon_{j,t}$$

where $log(N_{j,t})$ denotes the log-number of firms in industry i at time t; μ_j and η_t denote industry and time fixed effects; and $\Delta IP_{j,99-11} \times 1\{year\}$ denotes the interaction between Chinese import penetration from 1999 to 2011 and an indicator for each year. If Chinese competition leads to a reduction in the number of firms, we should find stable coefficients on the interaction term (β_t) before 2000; and decreasing coefficients thereafter. Figure 23 shows the results, which support our hypothesis. Chinese competition appears to have led to a statistically significant reduction in the number of firms.

Is the decline in the number of firms due to lower entry or higher exit? As shown in Figures 24 and 25, primarily lower entry. High exposure industries had traditionally higher entry rates than low exposure industries. But this pattern flipped in the early 2000's. Entry into high-exposure industries decreased and has remained well-below entry into low-exposure industries since 2003. By contrast, entry into low-exposure industries appears to have remained stable – affected primarily by the business cycle.

Figure 25 shows the 3-year moving average aggregate exit rates, and the percent exit rate through M&A, by level of import exposure. The total exit rates appear roughly similar across segments.

Figure 23: Number of Firms vs. Import exposure



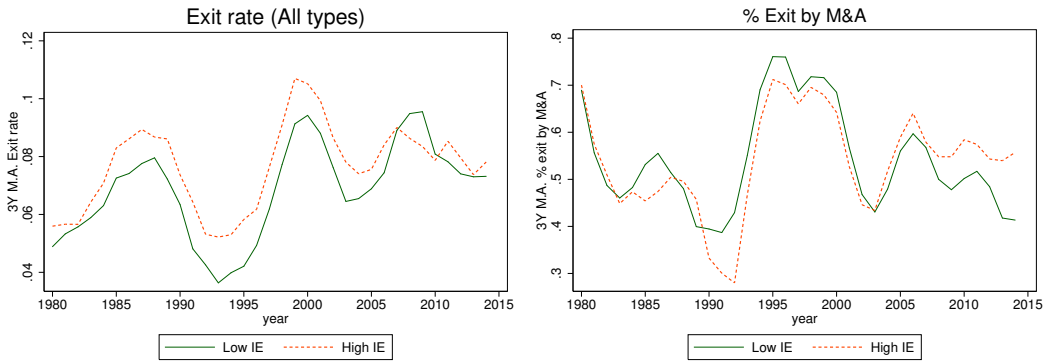
Notes: Figure shows the coefficients β_t from regressing $log(N_{j,t}) = \mu_j + \eta_t + \beta_t \Delta IP_{j,99-11} \times 1\{year\} + \varepsilon_{j,t}$. As shown, increased Chinese competition leads to a reduction in the number of firms. Annual data. Firm data from Compustat; import data from UN Comtrade. Includes only manufacturing industries.

Figure 24: Firm entry rate by Chinese exposure



Notes: Annual data. Firm data from Compustat; import data from UN Comtrade. Manufacturing industries only, split into 'high' (above-median) and 'low' (below-median) exposure based on import penetration from 1991 to 2011.

Figure 25: Firm exit rate and % exit through M&A, by Chinese exposure



Notes: Annual data. Firm data from Compustat; import data from UN Comtrade. Manufacturing industries only, split into 'high' (above-median) and 'low' (below-median) exposure based on import penetration from 1991 to 2011.

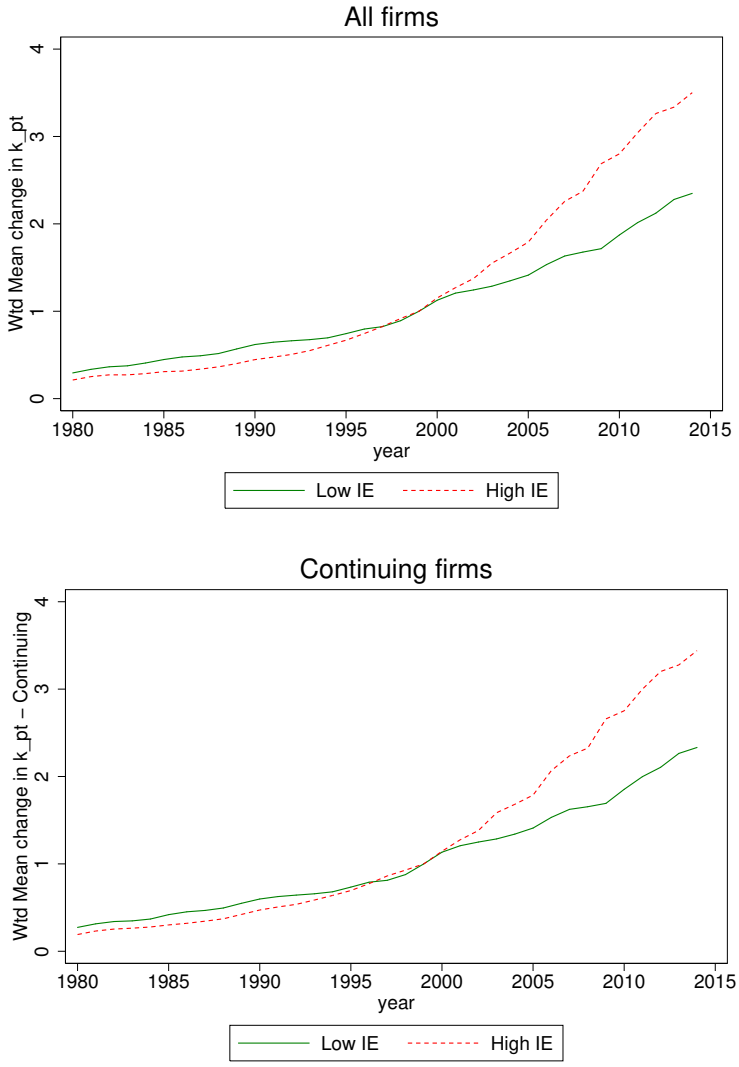
That said, exit through M&A increased for high exposure firms since the mid-2000's. Diving into industry-level exit rates also highlights some differences. In un-reported tests, we find that mean industry exit rate from 2000 to 2009 increases (significantly) with import exposure from 1991 to 2011. Thus, the substantially lower number of firms in high exposure industries appears to be primarily driven by lower entry, but also affected by higher exit and higher M&A activity.

Firm investment. Let us move on to investment. To more closely mirror the differences-in-differences structure of our main tests, we adjust the categorization of 'high' and 'low' exposure industries to those with above-median and below-median NTR Gap, respectively. Figure 9 in the main body already shows that the average firm that remained in Compustat following the China shock had a larger stock of K in high exposure industries than in low exposure industries. This Figure is created by dividing the aggregate stock of K in a given Compustat industry-year by the corresponding number of firms. It therefore accounts for firm entry and firm exit; and may not entirely represent the evolution of particular firms that lived through the China shock.

The top chart of Figure 26 addresses this by computing the weighted average change in total capital – relative to 1999 – for surviving firms in high and low exposure industries. Because we use 1999 as the benchmark year, only firms that were in Compustat as of that year are included in the calculation. The bottom chart takes the analysis a step further and considers only 'continuing' firms (i.e., firms that existed prior to 1995 and remain in the sample after 2009). Consistent with Figure 9, both plots show that firms in high exposure industries increased K faster than firms in low exposure industries.

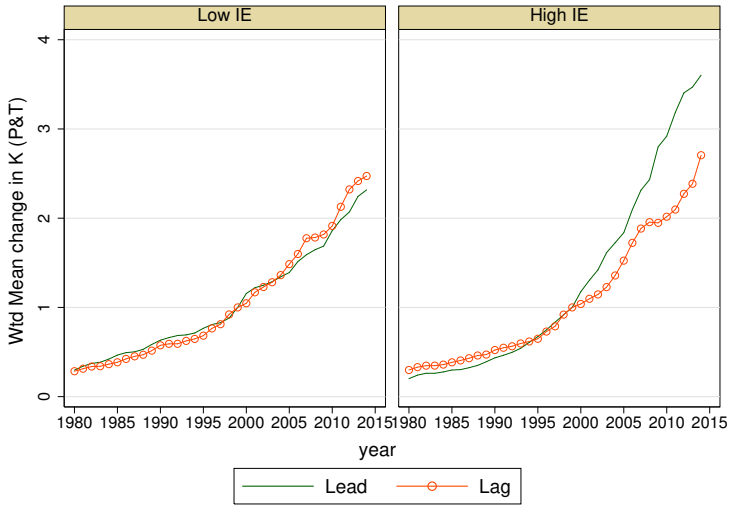
Moreover, the increase is concentrated in Leaders. Figure 27 replicates the top chart in Figure 26, but separating leaders and laggards within high and low exposure industries. In low exposure industries, leaders and laggards exhibit similar growth rates of capital. By contrast, leaders increase capital much faster than laggards in high exposure industries.

Figure 26: Log-change in K^{PT} by Chinese Exposure (1999 = 1)



Notes: Annual data. US incorporated firms in manufacturing industries only. Industries assigned to exposure based on NTR gap. Top plot includes all firms that existed as of 1999 and remain in the sample for a given year. Bottom plot includes only continuing firms (i.e., firms that existed prior to 1995 and remain in the sample as of 2009). Similar patterns for PP&E and Intangibles.

Figure 27: Log-change in K^{PT} for Leaders and Laggards, by Chinese Exposure



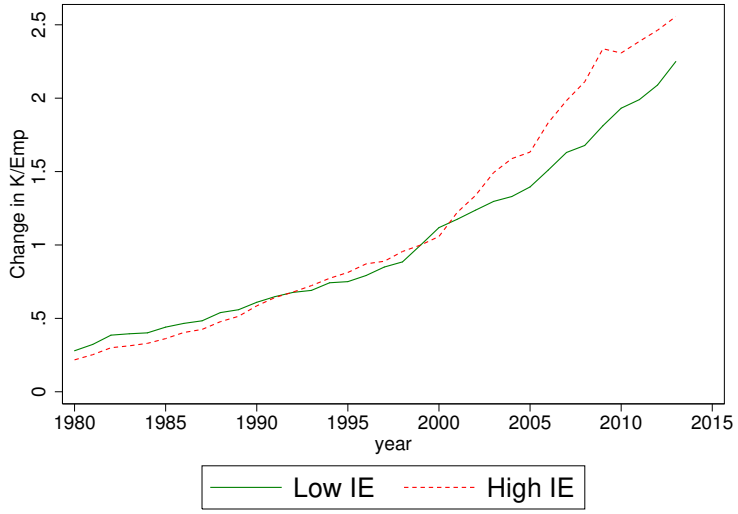
Notes: Annual data. US incorporated firms in manufacturing industries only. Industries assigned to exposure based on NTR gap. Leaders defined as firms with MV in top tercile as of 1999 within each NAICS Level 6 industry.

K/Emp. The effect of Chinese competition on employment has been widely studied – for instance in Pierce and Schott (2016); Acemoglu et al. (2016). They show that total employment decreased in industries most affected by Chinese competition. Pierce and Schott (2016) also show that K/Emp increased with Chinese exposure. We confirm these results using our Compustat sample. In particular, Figure 28 plots the evolution of K/Emp across high and low exposure firms. High exposure industries increase their K/Emp ratios, suggesting that they increased K faster than Emp (relative to low exposure industries).

Regression Results Last, we report several additional regression results excluding $\overline{\Delta IP_{j,t}^{US}}$ from our main specification to mirror Pierce and Schott (2016), and using$\Delta IP_{j,t}^{US}$ instrumented with import penetration of 8 other advanced economies $\left(\Delta IP_{j,t}^{OC}\right)$ to mirror Autor et al. (2016):

- Table 12: $log(K)$ results on $NTRGap$
- Table 13: $log(K)$ results on $\Delta IP_{j,t}^{US}$ instrumented by $\Delta IP_{j,t}^{OC}$
- Table 14: K/Emp results on $NTRGap_j \times \overline{\Delta IP_{j,t}^{US}}$
- Table 14: K/Emp results on $NTRGap_j$
- Table 16: K/Emp results on $\Delta IP_{j,t}^{US}$ instrumented by $\Delta IP_{j,t}^{OC}$

Figure 28: Change in K/Emp by Chinese Exposure (1999 = 1)



Notes: Annual data. Firm data from Compustat; import data from UN Comtrade. Manufacturing industries only, split into 'high' (above-median) and 'low' (below-median) exposure based on import penetration from 1991 to 2011.

Table 12: Chinese Competition: $\log(k_t)$ results based on $NTRGap_j$

Table shows the results of firm-level panel regressions of measures of capital on NTR gap. We consider three measures of capital: log-PP&E, log-intangibles and log-capital - where intangibles are measured as in Peters and Taylor (2016). Regression over 1991 - 2015 period. Leaders defined as firms with MV in top quartile as of 1999 within each NAICS Level 6 industry. Industry controls include measures of industry-level production structure (e.g.,PPE/Emp). As shown, leaders in industries with a higher NTR gap increased their capital relative to laggards after 2001. Annual data, primarily sourced from Compustat. Only US-incorporated firms in manufacturing industries included. Standard errors in brackets, clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01.

	(1) $log(PPE_t)^a$	(2) $log(Intan_t)^b$	(3) $log(k_t)^{a+b}$	(4) $log(PPE_t)^a$	(5) $log(Intan_t)^b$	(6) $log(k_t)^{a+b}$
$Post01 \times NTRGap$	-1.276**	-0.490+	-0.732**	-1.336**	-0.690**	-0.884**
	[0.32]	[0.26]	[0.22]	[0.41]	[0.26]	[0.26]
$Post01 \times NTRGap \times Lead_{99}$	**1.048****	**0.860****	**0.798****	**0.852****	**1.020****	**0.909****
	[0.27]	**[0.18]**	**[0.18]**	**[0.29]**	**[0.23]**	**[0.22]**
Observations	35830	35925	35936	17853	17832	17843
R^2	0.099	0.537	0.492	0.136	0.566	0.525
Industry controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Sample	All firms			Continuing firms		

Table 13: Chinese Competition: $\log(k_t)$ results based on $\Delta IP^{US}_{j,t}$ instrumented by $\Delta IP^{OC}_{j,t}$

Table shows the results of firm-level panel regressions of measures of capital on US-based import penetration, instrumented by import penetration at 8 other advanced economies. We consider three measures of capital: log-PP&E, log-intangibles and log-capital - where intangibles are measured as in Peters and Taylor (2016). Regression over 1991 - 2015 period. Leaders defined as firms with MV in top quartile as of 1999 within each NAICS Level 6 industry. Industry controls include measures of industry-level production structure (e.g.,PPE/Emp). As shown, leaders in industries with a higher NTR gap increased their capital relative to laggards after 2001. Annual data, primarily sourced from Compustat. Only US-incorporated firms in manufacturing industries included. Standard errors in brackets, clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01.

	(1) $log(PPE_t)^a$	(2) $log(Intan_t)^b$	(3) $log(k_t)^{a+b}$	(4) $log(PPE_t)^a$	(5) $log(Intan_t)^b$	(6) $log(k_t)^{a+b}$
$\widehat{\Delta IP^{US}_{j,t}}$	-0.435*	-0.505**	-0.449*	-0.409+	-0.580**	-0.542**
	[0.19]	[0.19]	[0.19]	[0.22]	[0.21]	[0.21]
$\Delta IP^{US}_{j,t} \times Lead_{99}$	**0.801****	**0.963****	**0.894****	**0.906***	**1.400****	**1.254****
	[0.28]	**[0.30]**	**[0.25]**	**[0.40]**	**[0.46]**	**[0.39]**
Observations	31311	31436	31441	14655	14674	14688
Within R^2	0.097	0.511	0.464	0.144	0.544	0.502
Industry controls	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES	YES	YES
Sample	All firms			Continuing firms		

Table 14: Chinese Competition: $\log(\frac{k_t}{Emp_t})$ results based on $NTR\,Gap_j \times \overline{\Delta IP^{US}_{j,t}}$

Table shows the results of firm-level panel regressions of measures of capital, employment and capital-deepening $NTR\,Gap_j \times \overline{\Delta IP^{US}_{j,t}}$, following 2. Regression over 1991 - 2015 period. Leaders defined as firms with MV in top quartile as of 1999 within each NAICS Level 6 industry. Industry controls include measures of industry-level production structure (e.g.,PPE/Emp). As shown, leaders increased capital, employment and K/Emp with the NTR Gap. Annual data, primarily sourced from Compustat. Only US-incorporated firms in manufacturing industries included. Standard errors in brackets, clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01.

	(1) $\log(k_t)$	(2) $\log(Emp_t)$	(3) $\log(\frac{k_t}{Emp_t})$
$Post01 \times NTR\,Gap$	-1.223	-2.927+	1.488
	[1.61]	[1.76]	[1.75]
$Post01 \times NTR\,Gap \times Lead_{99}$	**5.795****	**4.961****	**0.774**
	[1.33]	**[1.34]**	**[0.60]**
Observations	29982	29401	29380
Within R^2	0.46	0.053	0.365
Overall R^2	0.29	0.284	0.342
Industry controls	YES	YES	YES
Year FE	YES	YES	YES
Firm FE	YES	YES	YES
Sample	All firms		

Table 15: Chinese Competition: $\log(\frac{k_t}{Emp_t})$ results based on $NTRGap_j$

Table shows the results of firm-level panel regressions of measures of capital, employment and capital-deepening on NTR gap. Regression over 1991 - 2015 period. Leaders defined as firms with MV in top quartile as of 1999 within each NAICS Level 6 industry. Industry controls include measures of industry-level production structure (e.g.,PPE/Emp). As shown, leaders increased capital, employment and K/Emp with the NTR Gap. Annual data, primarily sourced from Compustat. Only US-incorporated firms in manufacturing industries included. Standard errors in brackets, clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01. 01.

	(1) $\log(k_t)$	(2) $\log(Emp_t)$	(3) $\log(\frac{k_t}{Emp_t})$
$Post01 \times NTRGap$	-0.732**	-0.544*	-0.205
	[0.22]	[0.26]	[0.22]
$Post01 \times NTRGap \times Lead_{99}$	**0.798****	**0.723****	**0.075**
	[0.18]	**[0.17]**	**[0.09]**
Observations	35936	35208	35107
Within R^2	0.492	0.069	0.405
Overall R^2	0.089	0.021	0.204
Industry controls	YES	YES	YES
Year FE	YES	YES	YES
Firm FE	YES	YES	YES
Sample	All firms		

Table 16: Chinese Competition: $\log(\frac{k_t}{Emp_t})$ results based on $\Delta IP^{US}_{j,t}$ instrumented by $\Delta IP^{OC}_{j,t}$

Table shows the results of firm-level panel regressions of measures of capital, employment and capital-deepening on US-based import penetration, instrumented by import penetration at 8 other advanced economies. Regression over 1991 - 2015 period. Leaders defined as firms with MV in top quartile as of 1999 within each NAICS Level 6 industry. Industry controls include measures of industry-level production structure (e.g.,PPE/Emp). As shown, leaders increased capital, employment and k/Emp with the NTR Gap. Annual data, primarily sourced from Compustat. Only US-incorporated firms in manufacturing industries included. Standard errors in brackets, clustered at the industry-level. + p<0.10, * p<0.05, ** p<.01.

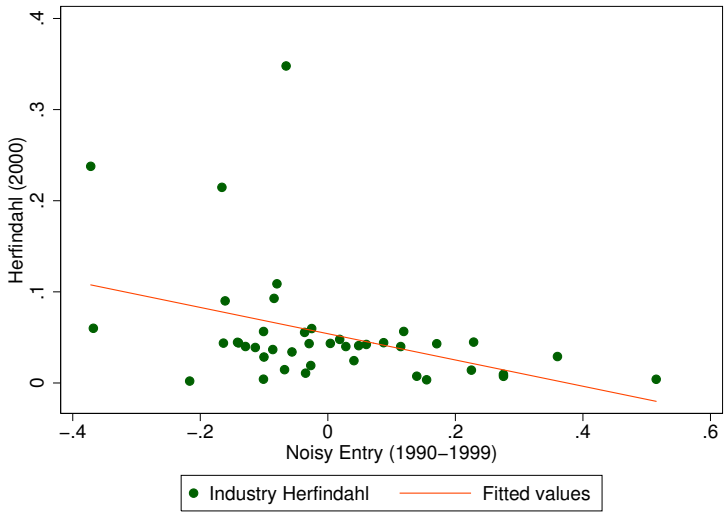
	(1) $\log(K_t)$	(2) $\log(Emp_t)$	(3) $\log(\frac{k_t}{Emp_t})$
$\widehat{IP^{US}_{j,t}}$	-0.449*	-0.227+	-0.233
	[0.19]	[0.13]	[0.20]
$\widehat{IP^{US}_{j,t} \times Lead_{99}}$	**0.894****	**0.738****	**0.150***
	[0.25]	**[0.26]**	**[0.07]**
Observations	31441	30815	30791
Within R^2	0.464	0.056	0.373
Industry controls	YES	YES	YES
Year FE	YES	YES	YES
Firm FE	YES	YES	YES
Sample	All firms		

Table 17: Post-2000 Entry and Exit vs. Pre-2000 Noisy entry: Regression Results

Table shows the results of industry-level OLS regressions of entry and exit measures on noisy entry. Entry and Exit based on the number of firms in Compustat. Standard errors in brackets. + p<0.10, * p<0.05, ** p<.01.

	(1) $\Delta logN_{00-05}$	(2) $\overline{Entry_{00-05}}$	(3) $\overline{Exit_{00-05}}$	(4) $\overline{M\&A\ Exit_{00-05}}$
$Noisy\ Entry_{i,90-99}$	**-0.366***	0.009	**0.049***	**0.039***
	[0.16]	[0.01]	**[0.02]**	**[0.02]**
$\overline{MedQ_{j,00,04}}$	0.12	0.031**	0.008	0.018*
	[0.09]	[0.01]	[0.01]	[0.01]
Observations	42	42	42	42
R-squared	0.173	0.31	0.145	0.212

Figure 29: Noisy entry (1990-1999) vs. Herfindahl (2000)



Notes: Annual data. See text for details.

C.2.2 Noisy Entry

This section presents additional exploratory analyses around our measure of noisy entry.

As a starting point, Table 17 formalizes the observations from Figure 10. It shows the results of regressing post-2000 changes in the number of firms, as well as average entry and exit rates, on pre-2000 noisy entry, controlling for Q. Higher noisy entry predicts a reduction in the number of firms, primarily due to higher exit.

Figure 29 shows that industries that experienced higher noisy entry in the 1990's had a lower import-adjusted Herfindahl in 2000. This is essentially the first stage of our regression, excluding the additional controls.

Table 18: Noisy Entry vs. Sales and Productivity: Regression Results

Table shows the results of industry-level OLS regressions of sales and value added on total and noisy entry. Sales and value added from BEA. TFP from U.S. KLEMS. Standard errors in brackets. + p<0.10, * p<0.05, ** p<.01.

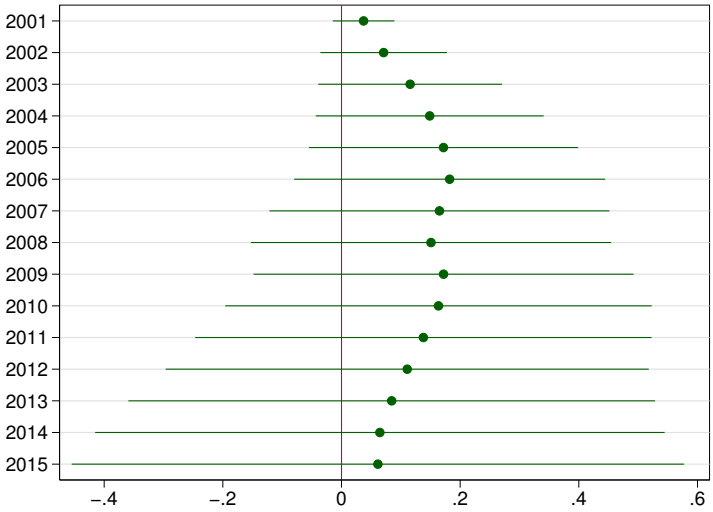
	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta log(Sale)_{99-04}$		$\Delta log(V.Add)_{99-04}$		ΔTFP_{99-04}	
$\Delta log(\#firms)_{94-99}$	0.019		**0.493****		**0.244****	
	[0.15]		**[0.15]**		**[0.07]**	
$Noisy\ Entry_{90-99}(i)$		-0.315		-0.26		-0.045
		[0.22]		[0.26]		[0.12]
Observations	43	42	43	42	43	42
R^2	0	0.05	0.208	0.025	0.256	0.004

Table 18 shows the results of regressing post-1999 changes in industry sales, value added and productivity on noisy entry. Actual entry predicts changes in value added and productivity but noisy entry does not. In fact, the coefficient shows the wrong sign.

Combined, the above results suggest the existence of substantial cross-sectional variations in noisy entry during the 1990's, which does not predict future demand or productivity. Noisy entry does, however, predict lower concentration in 2000; which makes it a valid instrument for industry-level investment.

Last, Figure 30 replicates the right plot of Figure 11 but considering $log\left(\frac{K_t}{K_{00}}\right)$ instead of NI/K . The results are consistent, with a rise in K immediately following the period of noisy entry, which reverts by the end of the sample. (Unreported) regression results yield similar conclusions.

Figure 30: Noisy entry coefficient on $log(K)$



Notes: Figure plots the coefficient of separate, annual regressions of changes in capital stock on our measure of noisy entry, following equation 5. As shown, industries with higher noisy entry experience a temporary increase in investment and capital. 10% confidence intervals are shown.

Table 19: Investment by Leaders: Noisy Entry

Table shows the results of firm-level panel regressions of the log change in the stock of capital (deflated to 2009 prices) on noisy entry. Regression from 2000 to 2015, following equation (1), except that HHI is replaced with $Noisy\,Entry_{j,90-99}$. We consider three measures of capital: PP&E, intangibles defined as in Peters and Taylor (2016) and their sum (total). Leaders measured as an indicator for firms having above median market value in the corresponding BEA segment k as of 2000. Q and log-age included as controls. As shown, leaders decrease investment with concentration, rather than increase it. Annual data, primarily sourced from Compustat. Standard errors in brackets, clustered at the firm-level. + p<0.10, * p<0.05, ** p<.01.

	(1)	(2)	(3)
	$\Delta log(PPE)^a$	$\Delta log(Int)^b$	$\Delta log(K)^{a+b}$
	$00-07$	$00-07$	$00-07$
Q(t-1)	5.349**	3.056**	3.733**
	[0.26]	[0.13]	[0.14]
$Noisy\,Entry_{90-99}(i)$	-13.457*	-10.804*	-17.536**
	[6.39]	[4.42]	[4.46]
$Noisy\,Entry_{90-99}(i) \times Lead_{99}$	**17.610***	**7.405**	**14.164****
	[6.95]	**[4.80]**	**[4.79]**
$Lead_{99}$	-10.585**	-5.477**	-6.718**
	[1.03]	[0.61]	[0.62]
$\log(Age_{t-1})$	-0.008	-6.022**	-5.401**
	[0.58]	[0.35]	[0.37]
Observations	29212	29348	29494
R^2	0.046	0.051	0.065
Year FE	YES	YES	YES
Firm FE	NO	NO	NO

Table 19 replicates the results of Table 4 but using noisy entry as a proxy for Competition instead of import-adjusted Herfindahls. It shows that leaders increase investment with noisy entry, consistent with our manufacturing results as well as the interaction of HHI and leaders.

Last, we consider an alternative specification to noisy entry based on explicit proxies of barriers to entry. In particular, we follow Frésard and Valta (2015) in conjecturing that deterring entry is less costly– and hence more likely – when rivals face more barriers to entry. We proxy barriers to entry with the number of citation-weighted patents as of 1995;[49] and estimate:

$$\begin{aligned}\Delta log(N)_{95,00} &= \beta_0 + \beta_1 log(CW\,Pat_{j,95}) + \beta_2 Excess\,K_{96,00} + \beta_3 Med\,Q_{j,96-00} + \beta_4 Med\,\Delta log\,Sales_{j,96-00} \\ &+ \beta_5 OS/K_{j,96-00} + \beta_6 CF/Assets_{j,96-00} + \beta_7 Med\,EPS\,Fcst_{j,00} \\ &+ \beta_8 \Delta IP^{US}_{j,91,99} + \beta_9 Mean\,firm\,assets_{90} + \beta_{10} Mean\,firm\,age_{90} + \varepsilon_j \\ log\left(\frac{K_{05}}{K_{00}}\right) &= \gamma_0 + \gamma_1 \widehat{\Delta log(N)}_{95,00} + \gamma_2 Excess\,K_{96,00} + \dots\end{aligned}$$

where we include as controls the same predictor variables as when estimating noisy entry. Thus, patents serve as an instrument for entry, above and beyond observables. The identification assump-

[49] Patent data is sourced from Kogan et al. (2017).

Table 20: Patents as an IV for Entry

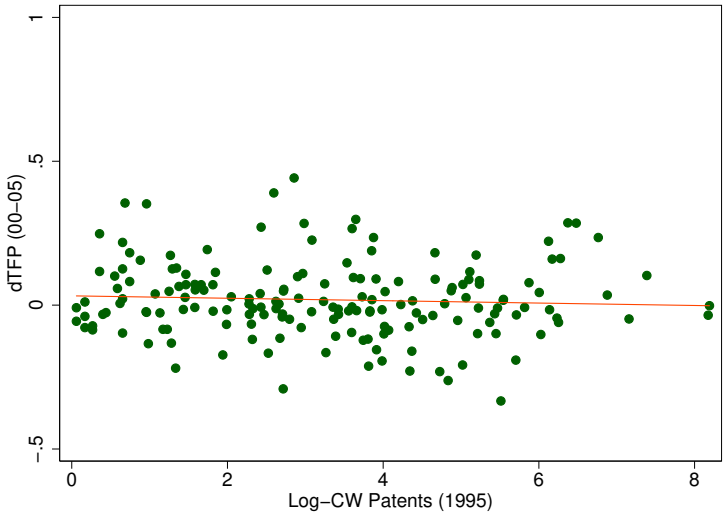
Table shows the results of industry-level panel regressions of changes in K and Net I/K on firm entry, instrumented by citation-weighted patent holdings. Controls include all variables used to estimate noisy entry. As shown, patent holdings led to lower levels of entry, which in turn predicted lower investment. Annual data, primarily from BEA. Standard errors in brackets, clustered at the firm-level. + p<0.10, * p<0.05, ** p<.01.

	(1)	(2)	(3)	(4)
	1st St.	2nd St.	1st St.	2nd St.
	$\Delta logN_{j,95-00}$	$log\left(\frac{K_{j,05}}{K_{j,00}}\right)$	$\Delta logN_{j,95-00}$	Net I/K
	00-05		00-05	
$\overline{mean\, Q_{j,95,00}}$	0.25**	-0.06		-0.024
	[0.09]	[0.07]		[0.02]
Mean Q (t-1)			-0.03	0.011
			[0.05]	[0.01]
$Ex.\, K_{j,95-99}$	2.20+	0.441		
	[1.10]	[0.78]		
$Excess\, Inv_{95-99}(i)$			2.43*	0.08
			[1.13]	[0.19]
$Log(CW\ Patents)_{j,95}$	**-0.03***		**-0.03****	
	[0.01]		**[0.01]**	
$\Delta logN_{j,95-00}$		**0.462+**		**0.097+**
		[0.24]		**[0.06]**
Other controls 95-00 (Sales,...)	...	...	...	...
Year FE	No		Yes	
Industry FE	No		No	
Observations	37	37	185	185
Within R^2		0.819		0.677
F-stat		4.26		10.60

tion here is that – controlling for other observables – patent-holdings as of 1995 are exogenous to expected demand and supply shocks that drive entry. The results are shown in Table 20. Higher patent holdings predict lower entry; and more entry predicts more investment.

The main identification concern would be that industries with high patent holdings as of 1995 experience positive productivity shocks following 2000. We test this directly using manufacturing industries, as shown in Figure 31. We find virtually no relationship between TFP growth and patent holdings.

Figure 31: ΔTFP vs. $log(CW\ Pat_{j,95})$



Notes: Figure plots the change in 5-factor TFP from 2000 to 2005 against log-transformed industry-level citation-weighted patent holdings as of 1995. TFP data from NBER-CES database covers NAICS Level 6 manufacturing industries. Patent holding data from Kogan et al. (2017)

D Model of Competition and Investment

The main contribution of our paper is empirical, but we need a model to understand precisely what the endogeneity issue is, and exactly what a valid experiment or a valid instrument would be. We use here a simple model of industry equilibrium under monopolistic competition. In addition to clarifying the identification problem, the model illustrates the role of information and suggests a set of potential instruments.

Description of the Model. The model is basically the industry block of a standard macroeconomic model. Firms make entry, investment, and production decisions. The timing is as follows:

- Period $t-1$: pay fixed cost κ^e_{t-1} to become active (or not)
- Period t: active firms are indexed by $i \in [0, N_{t-1}]$
  - Invest $k_{i,t}$;
  - Produce $y_{i,t} = A_t k_{i,t}^{\alpha} l_{i,t}^{1-\alpha}$;

All active firms have the same production function where A_t is productivity and $l_{i,t}$ is the quantity of labor hired by the firm. In terms of interpretation, it is best to think of each period as a few years. We assume that "firms" live for one period, or, equivalently, that the fixed cost κ must be repaid at the end of each period if the firm wants to remain active. The industry demand curve is given by the schedule

$$Y_t^D = D_t P_t^{-\sigma} \tag{9}$$

where D_t is a (stochastic) demand shifter, P_t is the industry price index, and σ is the demand elasticity across industries, which we assume is weakly above unity: $\sigma \geq 1$. Industry output is a CES aggregate of firms' outputs

$$Y_t^S \equiv \left(\int_0^{N_{t-1}} y_{i,t}^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}}, \tag{10}$$

where ϵ is the elasticity of substitution across firms within the same industry. It is larger than the between-industry elasticity: $\epsilon > \sigma \geq 1$.[50] The standard Dixit-Stiglitz CES aggregator takes ϵ as an exogenous parameter. It is straightforward to consider a model where ϵ is an increasing function of N_{t-1}, as in Feenstra (2003) for instance. This only reinforces the consequences of entry that we analyze below. For ease of exposition we treat ϵ as a parameter, and we simply point out in the discussion where endogenous firm-level markups matter. The price index is then defined in the usual way as

$$P_t \equiv \left(\int_0^{N_{t-1}} p_{i,t}^{1-\epsilon} di \right)^{\frac{1}{1-\epsilon}}, \tag{11}$$

[50] A standard calibration of the New Keynesian model is $\epsilon = 6$, chosen to deliver a steady state markup of 20% (Gali, 2008). The value of σ depends on the level of aggregation. In models with two sectors, home and foreign for instance, it is typical to use σ close to 1. With finer industry definitions σ should be higher, as assumed here.

and we impose the market clearing condition $Y_t^D = Y_t^S$.

The Firm's Problem. Let ρ_t be the user cost of capital at time t (i.e., the depreciation rate plus the rate of time preference).[51] The firm's problem is

$$\begin{aligned} &\max_{k_{i,t}, l_{i,t}, p_{i,t}} p_{i,t} y_{i,t} - w_t l_{i,t} - \rho_t k_{i,t}, \\ &\text{s.t. } y_{i,t} = \left(\frac{p_{i,t}}{P_t}\right)^{-\epsilon} Y_t, \\ &\text{and } y_{i,t} = A_t k_{i,t}^{\alpha} l_{i,t}^{1-\alpha}. \end{aligned}$$

The solution to the pricing problem is to set a fixed markup over marginal cost $p_{i,t} = \mu \chi_t$, where the markup is $\mu \equiv \frac{\epsilon}{\epsilon - 1}$ and the marginal cost is $\chi_t \equiv \frac{1}{A_t} \left(\frac{\rho_t}{\alpha}\right)^{\alpha} \left(\frac{w_t}{1-\alpha}\right)^{1-\alpha}$. In equilibrium, all firms set the same price and have the same size; which yields the industry price index (11) is $P_t = \mu \chi_t N_{t-1}^{\frac{-1}{\epsilon-1}}$.

The key point is that a larger number of firms implies a lower price index. The quantity index (10) becomes $Y_t^S = y_t N_{t-1}^{\frac{\epsilon}{\epsilon-1}}$. It increases with average firm output y_t, and with the number of firms because of the taste for variety. The industry equilibrium condition $Y_t^D = Y_t^S$ then implies $y_t = (\mu \chi_t)^{-\sigma} D_t N_{t-1}^{\frac{\sigma-\epsilon}{\epsilon-1}}$. Since $\epsilon > \sigma$, firm output is a decreasing function of the number of active firms, conditional on the demand shifter D_t. Investment is proportional to output: $k_t = \frac{\alpha}{\rho_t} \chi_t y_t$. Aggregating across firms we have $K_t = N_{t-1} k_t$, so industry investment is

$$K_t = \alpha \mu^{-\sigma} \frac{\chi_t^{1-\sigma}}{\rho_t} D_t N_{t-1}^{\frac{\sigma-1}{\epsilon-1}}. \tag{12}$$

We can summarize our results as follows:

Lemma 1. *Investment per firm and total industry investment both increase with demand D_t and decrease with the user cost ρ_t. Industry investment increases with the number of firms N_{t-1}, while investment per firm decreases.*

All firms have the same market share $\frac{1}{N_{t-1}}$ so the Herfindahl is predetermined and equal to the inverse of the number of firms:

$$H_t = \sum_1^{N_{t-1}} \left(\frac{1}{N_{t-1}}\right)^2 = \frac{1}{N_{t-1}} \tag{13}$$

The last step is to consider the entry decision of firms at time $t-1$.

[51] Formally, the firm's problem has two stages. At the production stage, the firm solves $\pi_{i,t}(k_{i,t}) \equiv \max_{l_{i,t}, p_{i,t}} p_{i,t} y_{i,t} - w_t l_{i,t}$, subject to the production function and the demand curve $y_{i,t} = \left(p_{i,t}/P_t\right)^{-\epsilon} Y_t$. At the investment stage it solves $\max_{k_{i,t}} \mathbb{E}_t \left[\pi_{i,t}(k_{i,t})\right] - \rho_t k_{i,t}$. Assuming that D_t is known at the beginning of time t, when $k_{i,t}$ is chosen, we can collapse the two stages into one.

Entry Decisions. The free entry condition is $\mathbb{E}_{t-1}\left[(\mu-1)\,\chi_t y_t\right] = (1+r_{t-1})\,\kappa^e_{t-1}$ where r_{t-1} is the required return on entry costs, taking into account the risk of failure as well as risk premia. Expected profits have to cover the entry cost. Note that the free entry condition essentially pins down expected firm output.[52] This is a typical property of expanding variety models. This free entry condition together with the equilibrium production derived earlier, pins down the number of firms:

$$N_{t-1}^{\frac{\epsilon-\sigma}{\epsilon-1}} = \frac{(\mu-1)\,\mu^{-\sigma}}{\kappa^e}\frac{\mathbb{E}_{t-1}\left[\chi_t^{1-\sigma}D_t\right]}{1+r_{t-1}}. \tag{14}$$

The number of active firms depends on expected demand, expected productivity (as long as $\sigma > 1$), and the cost of creating a firm. We now discuss the endogeneity issue. To make the point as clearly as possible, consider the following definition of a competitive economy.

Definition 1. *The competitive limit (denoted by c) with finite entry corresponds to $\epsilon \to \infty$ and $\kappa^e \to 0$, holding constant the ratio $\psi \equiv \frac{\mu-1}{\kappa^e}$.*

In the competitive limit, the markup μ converges to 1, and the entry cost converges to zero.[53] In the limit, we see from equation (12) that industry investment is independent from the number of firms

$$K_t^c = \frac{\alpha\chi_t^{1-\sigma}}{\rho_t}D_t, \tag{15}$$

and the industry price level equals the marginal cost, $P_t^c = \chi_t$, also independently from N_{t-1}.

Identification Problem. The competitive limit allows us to illustrate the endogeneity issue. By definition, market power is irrelevant in the competitive limit. Yet we will show that a regression of investment on standard measures of concentration would likely produce negative coefficients. To be more precise, consider an economy with competitive industries indexed by $j = 1..J$ subject to industry-specific demand shocks $D_{j,t}$. The aggregate economy is non-stochastic, and factors prices – w_t, ρ_t, and thus χ_t – are also non-stochastic. Investment in industry j is determined by equation (15) as $K_{j,t}^c = \frac{\alpha\chi_t^{1-\sigma}}{\rho_t}D_{j,t}$ and we specify the random demand shocks as

$$D_{j,t} = \bar{D}e^{d_{j,t-1}}e^{\nu_{j,t}}$$

where $\mathbb{E}_{t-1}\left[e^{\nu_{j,t}}\right] = 1$, and $d_{j,t-1}$ is known at time $t-1$ and has strictly positive cross-sectional variance: $VAR^{(j)}\left(d_{j,t-1}\right) > 0$. Suppose an econometrician runs cross-industry regressions (or panels regressions with time fixed effects) in order to determine the impact of concentration on investment. The proposition below explains the source of the bias.

Proposition 1. ***Fundamental Endogeneity Issue.*** *In the competitive limit with anticipated*

[52] So long as we abstract from the covariance between χ_t and y_t.

[53] What happens to the number of firms depends on the limit of the ratio ψ. A realistic benchmark is to have the ratio converge to a finite value. If the convergence of the markup to 1 is slower than that of the entry cost to 0 then N goes to infinity. The results below do not depend on having a "finite" number of firms, but it makes the exposition a lot simpler.

demand shocks, the cross-industry OLS regression of log-investment on log-Herfindahl gives a slope of minus one.

Proof. In the competitive limit with deterministic factor prices we have $\log N^c_{j,t-1} = \log \frac{\psi \chi_t^{1-\sigma}}{1+r_{t-1}} + \log \mathbb{E}_{t-1}[D_{j,t}]$, therefore $\log H_{j,t} = -\log \frac{\psi \chi_t^{1-\sigma}}{1+r_{t-1}} - \log \bar{D} - d_{j,t-1}$. The cross sectional variance of the Herfindahl is simply $VAR^{(j)}(\log H_{j,t}) = VAR^{(j)}(d_{j,t-1})$ On the other hand, we have $\log K^c_{j,t} = \log \frac{\alpha \chi_t^{1-\sigma}}{\rho_t} + \log \bar{D} + d_{j,t-1} + \nu_{j,t}$ therefore, using the Herfindahl index, $\log K_{j,t} = -\log H_{j,t} + \eta_t$ where η_t is a time fixed effect. This implies $COV^{(j)}\left(\log K^c_{j,t}, \log H_{j,t}\right) = -VAR^{(j)}(\log H_{j,t})$ and therefore the OLS slope is -1. □

This is an extreme example of omitted variable bias. There is in fact no economic connection between the number of firms and industry-level investment, as we can see from equation (15), but the econometrician would recover a coefficient of -1. The R^2 would depend on the variance of unexpected demand shocks $\nu_{j,t}$. A similar issue arises if we consider industry-specific productivity shocks $A_{j,t}$.

Corollary 1. *Industry-specific productivity shocks $A_{j,t}$ creates biases similar to the ones highlighted in Proposition 1.*

With industry productivity shocks, we have $\log H_{j,t} = -\log\left(\mathbb{E}_{t-1}\left[\chi_{j,t}^{1-\sigma}\right]\right) + ...$, and predictable cross-sectional variation in industry-level marginal cost, $\chi_{j,t}$, leads to biases as long as $\sigma > 1$, which is the realistic case.

Instruments and Natural Experiments. There are two ways to avoid the omitted variable bias. One solution would be to control for the demand shifter D_t. The problem is that D_t is not observable. We can only measure nominal sales $P_t Y_t$, which depend on both supply and demand factors.[54]

The other solution is to use natural experiments and/or instruments. The model can help us think about potential experiments and instruments. A good instrument in our model is a shock that randomly changes the opportunity cost of entry across industries. Let us consider the general model as a system of equations, where, as above, $j = 1..J$ indexes the industry:

$$\log K_{j,t} = \log \alpha \frac{\chi_{j,t}^{1-\sigma}}{\rho_t} + \log D_{j,t} - \sigma \log \mu_j + \frac{\sigma - 1}{\epsilon_j - 1} \log N_{j,t-1}$$

$$\frac{\epsilon_j - \sigma}{\epsilon_j - 1} \log N_{t-1} = \log(\mu_j - 1)\mu_j^{-\sigma} + \log \mathbb{E}_{t-1}\left[\chi_{j,t}^{1-\sigma} D_t\right] - \log(1 + r_{t-1})\kappa^e_{j,t-1}$$

This system makes it clear that random shocks to the entry cost κ^e_j could be used as instruments. More formally, we can state the following proposition.

[54] The exception is when $\sigma = 1$ (i.e., log-preferences for consumers) where nominal sales are exogenous from industry-level supply shocks, but this is not an assumption we can defend empirically.

Proposition 2. *Variation in entry costs κ_j^e that are uncorrelated with future demand $D_{j,t}$ and productivity $A_{j,t}$ would be valid instruments to assess the impact of concentration on investment.*

Our use of noisy entry essentially argues that the peculiar dynamics of entry in the late 1990's offer such an instrument. In particular, we document large cross-sectional variation in entry rates across industries, that do not predict future demand or productivity, but are driven by the willingness of investors (venture capitalists, or market participants in general) to fund risky ventures.[55]

A good natural experiment would be a change in the number of firms that is independent from future demand. Following the literature, we argue that the increase in Chinese competition; and particularly the formal entry of China in the WTO provides such an experiment. It comes, however, with two important caveats: first, it affects only manufacturing, which raises issues of external validity; second, it is a foreign competition shock, so it is unclear which prediction we can test using data on domestic investment. Interpreting the China shock therefore requires a model with firm heterogeneity, strategic competition, and foreign entry.

Large M&A assumes a discrete shock to competition that is sufficiently sharp to be uncorrelated with (long-run) changes in demand $D_{j,t}$ and productivity $A_{j,t}$.

[55] It does not matter for us whether the exuberance of the late 1990's was rational or not. Perhaps there were Bayesian mistakes, perhaps there were overly-optimistic forecasts, perhaps there were bubbles driven by the option to re-sell to future optimistic investors as in Scheinkman and Xiong (2003). At the end of the day, all that really matters is that these factors created variation in entry rates across industries (say in 1999) that turn out to be orthogonal to future demand (say in 2005).