

**WORKING PAPER** · NO. 2023-20

# 100 Years of Rising Corporate Concentration

*Spencer Y. Kwon, Yueran Ma, and Kaspar Zimmermann*

FEBRUARY 2023

# 100 Years of Rising Corporate Concentration<sup>\*</sup>

Spencer Y. Kwon<sup>†</sup>

Yueran Ma<sup>‡</sup>

Kaspar Zimmermann<sup>§</sup>

February 2023

## Abstract

We collect data on the size distribution of all U.S. corporations for 100 years. We document that corporate concentration (e.g., asset share or sales share of top businesses) has increased persistently over the past century. Rising concentration was stronger in manufacturing and mining before the 1970s, and stronger in services, retail, and wholesale after the 1970s. Furthermore, rising concentration in an industry aligns with greater technological intensity and more fixed costs. Industries with higher increases in concentration also exhibit higher output growth. Among the leading hypotheses for rising concentration, stronger economies of scale appear consistent with the long-run trends.

---

<sup>\*</sup>We thank Emi Nakamura and four anonymous referees for insightful suggestions. We also thank Daron Acemoglu, Douglas Baird, Nico Crouzet, Chang-Tai Hsieh, Erik Hurst, Steve Kaplan, Harald Kraus, Dmitry Kuvshinov, Huiyu Li, Chen Lian, John Moore, Sendhil Mullainathan, Brent Neiman, Christina Patterson, Carolin Pflueger, Thomas Philippon, Raghu Rajan, Esteban Rossi-Hansberg, José Scheinkman, Moritz Schularick, Doug Skinner, Andrei Shleifer, Amir Sufi, Chad Syverson, John Van Reenen, Rob Vishny, Anthony Zhang, Eric Zwick, and seminar participants at Berkeley, Bonn, Cambridge Judge, Cheung Kong GSB, Chicago Booth, Columbia, Dartmouth Tuck, Economics Dynamics Working Group, Edinburgh, EFA, Frankfurt School of Finance and Management, ifo Conference on Macroeconomics and Survey Data, Hong Kong University, Kansas City Fed, Leibniz Institute for Financial Research SAFE, Michigan Ross, MIT Sloan, NBER EF&G Meeting, NBER Megafirms and the Post-COVID Economy Meeting, Philadelphia Fed, Pittsburgh Katz, PKU HSBC, Princeton, and PUC Chile for very helpful comments and suggestions. We are grateful to Fatin Alia Ali, Hadi Allahkhah, Christian Alshon, Daniel Zongsheng Huang, James Kiselik, Mohamed Mahmoud, Derek Hsieh, Koichi Onogi, Ramraj Harikanth Sundararaj, Arian Tabibian, Juan Uphoff, Julien Weber, Zephyr Xu, and Yingxuan Yang for excellent research assistance. We thank Mike Cusick at the BEA, Victoria Triplett at the National Archives, and staff members at the IRS for their kind help with data inquiries. First draft: August 2021.

<sup>†</sup>Harvard University (yongwookkwon@g.harvard.edu).

<sup>‡</sup>University of Chicago Booth School of Business and NBER (yueran.ma@chicagobooth.edu).

<sup>§</sup>Leibniz Institute for Financial Research SAFE (zimmermann@safe-frankfurt.de).

# 1 Introduction

The role of large businesses in the economy is an important question for researchers, policymakers, and the public. The finding of rising concentration among U.S. industries since the 1980s (as shown by [Autor et al. \(2020\)](#) and others using comprehensive census data covering this period) has attracted particular attention. Recent discussions of this evidence often focus on the special features of today's world. In the archives of history, however, lives an old conjecture that rising concentration is a feature, if not a law, of industrial development. In fact, both [Marx \(1867\)](#) and [Marshall \(1890\)](#) wrote that technological progress increases economies of scale and raises the concentration of production. [Lenin \(1916\)](#) gathered census statistics in the early 1900s to back up the conviction that “the enormous growth of industry and the remarkably rapid concentration of production...are one of the most characteristic features of capitalism.” After the National Bureau of Economic Research (NBER) was established in 1920 to provide “exact and impartial determinations of facts,” one of its first publications also noted the spread of mass production at that time, as well as a general view that “each generation believes itself to be on the verge of a new economic era” but what appears to be new often represents recurring themes in history ([Committee on Recent Economic Changes, 1929](#)).

In this paper, we collect data covering the population of U.S. corporations for 100 years, from 1918 to 2018. We use these data to study the long-run evolution of the concentration of production, namely the extent to which a small set of top businesses account for a large share of production assets or output. We document that corporate concentration in the U.S. (measured using the asset share, sales share, or net income share of top businesses) has increased over the past century. Among different sectors, rising concentration was stronger in manufacturing and mining before the 1970s, and stronger in services, retail, and wholesale after the 1970s. We then examine the leading hypotheses about the economic mechanisms behind rising concentration, including economies of scale, globalization, regulation, among others. Overall, our data show that increasing concentration of production activities has been a feature of the U.S. economy for at least a century, and long-run forces can be important for this development.

We obtain long-run data on the business size distribution in the U.S. by digitizing historical publications of the Statistics of Income (SOI) and the associated Corporation Source Book from the Internal Revenue Service (IRS). Since 1918, the SOI has been reporting annual statistics of the population of corporations by size bins, including the number of businesses and their financial information (e.g., assets, sales, net income). We use these size bins to estimate top businesses' shares in the aggregate, the main sectors (roughly one-digit SICs), and the subsectors (roughly two-digit SICs). Our baseline estimation method uses the generalized Pareto approach by fitting a Pareto curve to each size bin ([Blanchet, Fournier, and Piketty, 2022](#)); results are over 0.99 correlated if we instead fit a lognormal curve to each size bin, or directly add the top bins up to a given number of businesses. The SOI data capture production activities in the U.S. (similar to the gross output convention in the national accounts), which align with our focus on the concentration of production.

For the aggregate economy, the data reveal a persistent rise in the shares accounted by top businesses. In earlier years (from 1918 to 1975), the SOI provide size bins sorted by net income, and we use these data to compute the share of top corporations by net income in total corporate net income (restricting

to businesses with positive net income). In later years (from 1959), the SOI provide size bins by receipts (sales), and we use these data to compute the share of top corporations by receipts in total corporate receipts. The longest and most comprehensive size bin tabulations are sorted by assets (since 1931), and we use these data to compute the share of top corporations by assets in total corporate assets. The long-run increase in corporate concentration is reflected by all three series. For instance, since the early 1930s, the asset shares of the top 1% and top 0.1% corporations have increased by 27 percentage points (from 70% to 97%) and 40 percentage points (from 47% to 88%), respectively.

At the industry level, we also observe a general rise in corporate concentration among the main sectors (where all three types of size bins are available) and the subsectors (where only size bins by assets are available). However, the timing differs across industries. For manufacturing and mining, rising concentration was stronger in earlier decades (before the 1970s); for services, retail, and wholesale, rising concentration was stronger in later decades (after the 1970s). The results are similar if we examine the relative concentration within the largest businesses (e.g., the top 1% relative to the top 10%).<sup>1</sup> The overall trends are also similar if we use a fixed number of top businesses (e.g., top 500 or 5,000), but a fixed number can be less comparable across different levels of aggregation (e.g., economy as a whole versus a particular industry) or across different industries that vary substantially in size.

We perform several additional checks for the concentration trends. First, our main analyses use comprehensive size bins for corporations (both C- and S-corporations). For noncorporations (partnerships and nonfarm proprietorships), we can obtain size bins by receipts for some years. In these years, we construct the receipt share of top businesses by receipts among corporations plus noncorporations, and we find similar results of rising concentration. Second, we cross-check with census data, which report sales shares of the top 4, 8, 20, and 50 firms by sales in census years, for manufacturing industries since 1947 and other industries since the 1980s. For the degree of concentration over time, we rely on the longer manufacturing census data at the four-digit SIC level. We take the average concentration ratios in each census year and observe an upward trend. For the degree of concentration in the cross section, we use recent data from 2012 (with both manufacturing and non-manufacturing industries) and compare top shares in census data with those interpolated from SOI data. These two sets of data match in level and are 0.84 correlated. Third, we use Compustat data to check that our SOI data align with firms' financial statements and reliably capture firms at the very top. We calculate the total sales (assets) among the top 500 Compustat firms by sales (assets) as a share of total corporate sales (assets). Compustat data miss large private companies and have other caveats, but the level of aggregate top 500 shares is comparable with estimates from SOI data. Fourth, in recent years some U.S. firms may have shifted their production assets to foreign subsidiaries for various reasons. We show that the results are similar if we include the international assets of U.S. companies, using data from the BEA on Activities of U.S. Multinational Enterprises (available since the 1980s).

In addition to the business size distribution, the SOI also provides information about other firm characteristics over the past century, such as profitability. Profitability (i.e., net income/sales) does not display secular trends, unlike concentration. It plunged during the Great Depression, rebounded during

---

<sup>1</sup>A possible concern about the baseline top 1% share is that when small firms enter the number of firms increases, which may mechanically increase the top 1% share. The relative concentration measure (the top 1% within the top 10%) is robust to this concern as we discuss more in Section 2 and it is 98% correlated with our baseline top 1% share.

the 1940s, then declined gradually until the 1980s, and increased slightly afterwards.

Why did corporate concentration increase persistently over the past century? We examine the leading hypotheses about rising concentration, which have attracted considerable attention in recent research. It is inevitably challenging to pin down the exact cause, but some mechanisms appear more consistent with the long-run trends and the timing of rising concentration in different industries.<sup>2</sup> We begin with economies of scale, for which we observe a reasonable amount of supportive evidence. At a high level, the long-run concentration trends are consistent with industrial technologies enabling large-scale production in manufacturing in the early 20th century (Chandler, 1994), and modern IT transforming services, retail, and wholesale more recently (Hsieh and Rossi-Hansberg, 2022). We present three sets of results in line with the predictions of economies of scale.

First, we find that the timing and the degree of rising concentration in an industry align closely with rising technological intensity. For each industry, we obtain long-run annual data on the investment intensity of R&D and IT from the Bureau of Economic Analysis (BEA). These types of investment are commonly viewed to embody greater scalability (Haskel and Westlake, 2017; Crouzet and Eberly, 2019; Lashkari, Bauer, and Boussard, 2022) and entail greater upfront spending (Sutton, 1991, 2001). We observe that the top 1% share in an industry comoves strongly with the investment intensity of R&D and IT, using both levels and changes over the medium term (e.g., twenty years). We also include specifications with year fixed effects to absorb aggregate trends, which further isolate the timing alignment between rising concentration in an industry and its technological intensity. We then examine technological innovations in production processes measured using long-run data on breakthrough patents (Kelly et al., 2021), available mainly for manufacturing subsectors plus mining, construction, and agriculture. The patent data show that influential technologies are associated with more production concentration (whereas the total number of patents per se does not play a role).

Second, we find that the degree of concentration is positively correlated with a measure of the intensity of fixed operating costs. Several studies have been interested in detecting fixed operating costs using financial statements (Anderson, Banker, and Janakiraman, 2003; De Loecker, Eeckhout, and Unger, 2020; Traina, 2018), especially through the decomposition of operating costs into costs of goods sold (COGS, often thought to be more variable) and selling, general, and administrative expenses (SG&A, often thought to be more fixed). We allow a portion of each category to be variable, and estimate this portion by regressing log changes in COGS (SG&A) on log changes in sales among Compustat firms in each industry (for the average industry, COGS and SG&A change by 0.75% and 0.46% for a 1% change in sales, respectively). We add up the corresponding portion of COGS and SG&A to obtain total variable costs, normalize by total operating costs, and take the remainder to be the fixed cost share. We use the median fixed cost share for each industry in a given year among Compustat firms. Compustat firms are generally large and likely to utilize scalable technologies, so they can be useful for reflecting the relevance of scalable technologies with higher fixed costs. We observe that concentration comoves with the intensity of fixed operating costs, using both levels and changes within an industry over time.

---

<sup>2</sup>We have not focused on the random growth mechanism because standard random growth frameworks do not easily explain systematic differences across industries (Sutton, 1997). In addition, these frameworks generally focus on a stationary size distribution, while the empirical evidence suggests that the business size distribution is not necessarily stationary for a reasonably long period of time.

Third, we find that increases in concentration are positively associated with industry growth. Over the medium term, industries that experience higher increases in concentration are also the ones that experience higher growth in real gross output. Correspondingly, their output shares in the economy expand. We check that the results are robust when we use the persistent component of industry growth (e.g., predicted by industry developments in the past), which addresses the concern that random shocks to large firms can affect both changes in concentration and contemporaneous industry growth.

We use a simple model to illustrate that production processes with greater economies of scale can account for the empirical facts. Following [Hsieh and Rossi-Hansberg \(2022\)](#), firms with heterogeneous productivity can choose a new production technology with higher upfront spending and lower marginal costs, or an old technology with lower upfront spending and higher marginal costs. Productive firms will find it worthwhile to pay the higher upfront spending and obtain scalability; other firms will not find it appealing to do so, but they can still exist when products are imperfect substitutes. Greater scalability of the new technology will increase concentration (e.g., top 1% share) as large firms expand relative to small firms; industry output will also increase. Meanwhile, profitability can follow other forces that drive markups.

We then examine whether trade and globalization can explain our baseline facts. International trade for the U.S. (relative to GDP) did not expand in the first half of the 20th century, and globalization only started to accelerate around the 1970s. For manufacturing, rising concentration was stronger before the globalization era; for services, rising concentration has been stronger in recent decades, but the volume of international trade is smaller in services. The timing suggests that international trade does not account for the entire long-run evidence. Additionally, in a simple model following [Melitz \(2003\)](#), if the reduction of barriers to international trade is the only force, then the concentration of domestic sales (i.e., sales excluding exports) for U.S. companies would not increase. In the data, the concentration of sales excluding exports has increased significantly as well. Nonetheless, even though trade and globalization alone cannot fully explain the long-run evidence, having broad markets (domestically or internationally) could increase the appeal of economies of scale.

Another question is whether regulatory policies drive our main facts. For antitrust, the past century witnessed several enforcement regimes ([Lamoreaux, 2019](#)), whereas rising corporate concentration has been a persistent phenomenon. In the data, we do not observe a significant relationship between corporate concentration and standard aggregate antitrust enforcement measures, such as the number of antitrust cases filed by the Department of Justice (DOJ) or the budget of the DOJ's antitrust division. While we do not find evidence that antitrust shapes the economy-wide business size distribution, it could have a more visible impact on a particular market defined for antitrust analyses ([Affeldt et al., 2021](#)). For other regulations, restrictions on interstate banking directly affect the size of banks (and we indeed observe rising concentration in banking after these restrictions were lifted), and various types of government policies and programs may directly or indirectly influence the size of businesses. Overall, to explain our findings, regulations need to have shifted in favor of large firms over the past century; moreover, they need to have been particularly important for the expansion of large firms in manufacturing rather than in services in the earlier decades, and then switched focus in the later decades. At the moment, we are not aware of such patterns in regulatory policies.

Finally, changing demographics or low interest rates could affect the overall level of corporate concentration in recent decades, but they might not explain the long-run trends throughout the past century or the timing of rising concentration in different industries. Decreasing search frictions or higher elasticity of substitution for buyers could also increase concentration, but it is more challenging to measure these buyer-side changes systematically and less clear why such changes would have affected different industries at different points in time.

Even if rising production concentration follows from stronger economies of scale, the welfare implications can be nuanced. For example, [De Loecker, Eeckhout, and Mongey \(2022\)](#) point out that the welfare impact is ambiguous if fixed costs of production rise; in this case, the more efficient producers become more dominant, but more resources are tied up in overhead costs. [Aghion et al. \(2022\)](#) present a model where concentration increases due to lower costs of spanning multiple products or a rising efficiency advantage of large firms. However, welfare declines given their baseline parameter values because new innovators would be more likely to face a high-efficiency existing firm, which discourages innovation and growth in the longer term.

As reviewed by [Syverson \(2019\)](#), higher concentration can be associated with either less or more competition, and markups are better barometers for market power. Correspondingly, we do not speak to the strength of market power or the effectiveness of policies targeting market power; a number of other studies examine these issues in detail ([De Loecker, Eeckhout, and Unger, 2020](#); [Gutiérrez and Philippon, 2022](#)). Some postulate that economies of scale may increase market power ([Lenin, 1916](#); [Lange, 1937](#); [Persky, 1991](#); [Eeckhout, 2021](#); [Eeckhout and Veldkamp, 2022](#)). Estimated markups in the literature do not appear to increase before the 1980s; combined with our findings, the evidence suggests that stronger economies of scale may not always raise market power. Analyzing the conditions for economies of scale to increase market power is an interesting topic for future research.

Our work contributes to knowledge about the long-run evolution of the U.S. economy in three ways. First, researchers and the general public have been interested in the concentration of production and the role of large enterprises for a long time, as illustrated at the beginning. Several influential recent studies analyze industry concentration in the U.S. since the 1980s using census data ([Autor et al., 2020](#); [Covarrubias, Gutiérrez, and Philippon, 2020](#)). Several early studies analyze a few years of SOI data to calculate the shares of the top 100 or 200 corporations in the economy during the initial decades of the 20th century, which often point to an increasing pattern ([Means, 1931](#); [Means et al., 1939](#); [Blair, Houghton, and Rose, 1946](#); [Adelman, 1951](#); [Collins and Preston, 1961](#)). We offer comprehensive long-run evidence, and provide datasets for researchers to study the U.S. business size distribution throughout the past century. Relatedly, several papers present long-run evidence on the “size distribution” of household income and wealth, such as [Piketty and Saez \(2003\)](#) and [Saez and Zucman \(2016\)](#); we show that similar historical administrative data for businesses contain extensive information that helps us better understand heterogeneity among firms.<sup>3</sup>

---

<sup>3</sup>For analyses of recent decades using data on public companies, see also [DeAngelo, DeAngelo, and Skinner \(2004\)](#), [Grullon, Larkin, and Michaely \(2019\)](#), and [Gutiérrez and Philippon \(2020\)](#). Some analyses examine a small number of “giant” or “dominant” firms, such as top 20 in the aggregate economy, which tend to find weaker trends in the shares of this set of firms. We capture a broader set of top businesses, and the rise of corporate concentration seems not limited to a few giant companies. Consistent with the conjecture of [White \(2002\)](#), the increasing prevalence of larger enterprises in the right tail



Second, a rapidly growing body of research has proposed various economic mechanisms for rising concentration among U.S. firms, including economies of scale and technology (Ganapati, 2021; Hsieh and Rossi-Hansberg, 2022; Hubmer and Restrepo, 2022; Firooz, Liu, and Wang, 2022), regulation (Philippon, 2019; Akcigit and Ates, 2022; Singla, 2022), demographics (Hopenhayn, Neira, and Singhanian, 2022; Peters and Walsh, 2022), low interest rates (Kroen et al., 2022; Zhao, 2022), among others. Our findings show that a full account of rising concentration needs to explain not only the recent trends, but also the long-run facts. Establishing a given mechanism is inevitably challenging, but the long-run data appear to align most closely with stronger economies of scale. Meanwhile, each decade may have special features that are important to examine, and our long-run analysis complements studies that scrutinize particular decades.

Third, the size distribution of businesses is important for the determinants of macroeconomic outcomes: the features and the actions of large companies matter more with higher production concentration. For example, many studies analyze financial frictions among large versus small businesses; as Crouzet and Mehrotra (2020) point out, high and rising concentration implies that the frictions that apply to the bottom 99% of firms have a modest impact on aggregate fluctuations. Gabaix (2011) also highlights that shocks to large firms can drive aggregate fluctuations; such effects are likely stronger when production concentration is higher.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 presents the main facts about corporate concentration in the past 100 years and a number of robustness checks. Section 4 examines the leading hypotheses of rising concentration. Section 5 concludes.

## 2 Data

Our main data source is the Statistics of Income (SOI) and the associated Corporation Source Book published annually by the IRS. The SOI originated from the Revenue Act of 1916, which requires the IRS to report statistics based on the tax returns filed each year. Statistics on the size of corporations were included for the first time in 1918. The SOI is a key source for the national income and product accounts (NIPA), and the data we use capture production activities in the U.S. like the national accounts. For the early years (before 2000), we digitize historical SOI publications; for some years between 1965 and 1980, we are able to use data from the Electronic Records Archives of the National Archives. For recent years (after 2000), we use data from the SOI website. We explain our data construction in this section and provide more details in Internet Appendix IA2. We have made the cleaned data series available at <https://businessconcentration.com> and in our replication package.

Every year, the SOI tabulates a variety of statistics for the population of corporations, including the number and financial information of corporations by size bins, which allows us to investigate the business size distribution. Table 1 shows examples for the aggregate economy (Panel A) and for one industry (Panel B) from the SOI in 1945. The primary size bins that we use are based on total

---

of the size distribution can be related to changes in technology. Although economy-wide production concentration has been increasing, concentration in the market for a particular product or location may decrease as large firms expand into more markets (Rossi-Hansberg, Sarte, and Trachter, 2021; Benkard, Yurukoglu, and Zhang, 2021).



assets, since these size bins are reported continuously for the longest period of time and have the most detailed breakdowns by industry. Size bins are also available based on receipts (sales) after 1959 and net income from 1918 to the 1970s. These data on corporations by size include both C-corporations and S-corporations, and the information comes from corporate tax returns (various types of 1120 forms). For noncorporations (partnerships and nonfarm proprietorships), size bins by receipts are presented in some years in separate SOI publications. We transcribe these data whenever available, and we show in Section 3.1 that the top 1% receipt shares among corporations and among all businesses (corporations plus noncorporations) are similar.<sup>4</sup>

We have processed tabulations of corporations by size bins for the aggregate economy, main sectors (roughly at the one-digit SIC code level), and subsectors (roughly at the two-digit SIC code level). The industry classification system switched from SIC to NAICS in 1997, and we harmonize the industries to maintain consistency as explained in Internet Appendix IA2; the list of main sectors and subsectors (together with the mappings to SIC and NAICS codes) is shown in Table IA6. The SOI data cover all industries, unlike the census data for concentration ratios which exclude some sectors like agriculture. The SOI assigns a single industry code to each business based on the industry that represents the largest percentage of its total receipts.<sup>5</sup> The IRS stopped publishing the sector-level tabulations after 2013 due to an update in its privacy guidelines (IRS Publication 1075), but the aggregate tabulations continue to be available. Information at the individual firm level is confidential and not available in our data.

For each level of aggregation, we estimate top business shares from size bins marked by dollar thresholds. We verify that different estimation methods produce similar results in Figure IA1. First, we can fit certain distributions to the raw data. Our baseline method uses the generalized Pareto interpolation, which is the standard approach to estimate household top income shares from income bins with a similar format. Blanchet, Fournier, and Piketty (2022) provide a detailed description of this method, which refines and standardizes top share interpolations in earlier work (e.g., Piketty and Saez, 2003). The generalized Pareto interpolation starts by calculating the inverted Pareto coefficient  $b(p_i)$  for each bin threshold  $i$ , where  $p_i$  is the fraction of firms with assets (receipts/net income) more than  $y_i$ , and  $b(p_i)$  is the ratio between the average assets (receipts/net income) above  $y_i$  and the threshold  $y_i$ . It then derives a continuous curve of inverted Pareto coefficients. Alternatively, we can also fit a lognormal curve for each bin and interpolate the lognormal curves, as explained in Internet Appendix IA2.4. Some studies investigate whether Pareto or lognormal provides a close fit of the business size distribution when a given parameter is used for the entire distribution (Rossi-Hansberg and Wright, 2007; Kondo, Lewis, and Stella, 2019). Our interpolation allows the Pareto or lognormal parameters to vary for different parts of the distribution, in which case the top share using either method is similar. Pareto has one parameter to fit whereas lognormal has two, so the former is easier to implement.

Second, we can directly add up the top bins such that the number of businesses in these bins

---

<sup>4</sup>It is challenging to construct reliable economy-wide measures of corporate concentration before our SOI data started in 1918, since limited information exists to our knowledge about the denominator (total corporate assets or sales), the numerator (the largest businesses and their assets or sales), as well as the number of businesses.

<sup>5</sup>Since we are interested in the business size distribution, we do want to keep a business as a whole instead of separating it into different pieces. The SOI industry classification is in line with this objective (and ensuring that a business is not counted multiple times in different industries), although the industry classification may have some imperfections.

Table 1 – Raw Data from Statistics of Income (1945)

## Panel A. Example of Aggregate Tabulation

[Total assets classes and money figures in thousands of dollars]

| Total assets classes <sup>38</sup> | Number of returns <sup>32</sup> | Total as-sets—Total liabilities <sup>38</sup> | Total compiled receipts <sup>6</sup> | Compiled net profit or net loss | Net income or deficit <sup>2</sup> |
|------------------------------------|---------------------------------|-----------------------------------------------|--------------------------------------|---------------------------------|------------------------------------|
| AGGREGATE                          |                                 |                                               |                                      |                                 |                                    |
| Under 50.....                      | 177,788                         | 3,647,660                                     | 9,030,941                            | 267,783                         | 267,621                            |
| 50 under 100.....                  | 61,431                          | 4,378,846                                     | 8,650,707                            | 376,597                         | 376,379                            |
| 100 under 250.....                 | 60,308                          | 9,526,342                                     | 16,659,649                           | 837,872                         | 837,120                            |
| 250 under 500.....                 | 27,583                          | 9,666,507                                     | 15,828,823                           | 914,465                         | 913,563                            |
| 500 under 1,000.....               | 17,669                          | 12,436,856                                    | 17,397,634                           | 1,196,416                       | 1,193,741                          |
| 1,000 under 5,000.....             | 22,057                          | 47,907,402                                    | 42,250,752                           | 3,450,003                       | 3,427,380                          |
| 5,000 under 10,000.....            | 3,948                           | 27,591,330                                    | 17,749,140                           | 1,719,313                       | 1,704,217                          |
| 10,000 under 50,000.....           | 3,197                           | 65,334,850                                    | 39,917,400                           | 3,900,112                       | 3,868,073                          |
| 50,000 under 100,000.....          | 427                             | 29,834,282                                    | 15,626,460                           | 1,521,776                       | 1,508,085                          |
| 100,000 and over.....              | 542                             | 231,137,144                                   | 69,524,822                           | 7,035,344                       | 6,917,796                          |
| Total.....                         | 374,950                         | 441,461,268                                   | 252,636,330                          | 21,219,681                      | 21,013,975                         |

## Panel B. Example of Industry-Level Tabulation

| Total assets classes <sup>55</sup> | Num-ber of returns with balance sheets <sup>42</sup> | Cash <sup>43</sup> | Notes and accounts receivable less reserve | Inven-tories | Invest-ments <sup>50</sup> | Capital assets <sup>46</sup> less reserves | Total assets—Total lia-bilities <sup>48</sup> | Accounts and notes payable <sup>57</sup> |
|------------------------------------|------------------------------------------------------|--------------------|--------------------------------------------|--------------|----------------------------|--------------------------------------------|-----------------------------------------------|------------------------------------------|
| SERVICE: BUSINESS SERVICE—         |                                                      |                    |                                            |              |                            |                                            |                                               |                                          |
| 0.....                             | 2,419                                                | 11,723             | 11,877                                     | 1,264        | 4,076                      | 9,901                                      | 41,710                                        | 9,258                                    |
| 50.....                            | 519                                                  | 9,009              | 11,138                                     | 1,170        | 4,203                      | 8,186                                      | 36,246                                        | 8,416                                    |
| 100.....                           | 433                                                  | 13,848             | 20,844                                     | 1,964        | 10,960                     | 15,292                                     | 66,950                                        | 15,679                                   |
| 250.....                           | 173                                                  | 12,547             | 16,526                                     | 2,122        | 11,386                     | 13,542                                     | 60,054                                        | 15,626                                   |
| 500.....                           | 99                                                   | 11,875             | 20,417                                     | 2,614        | 16,166                     | 11,856                                     | 66,486                                        | 17,827                                   |
| 1,000.....                         | 92                                                   | 31,130             | 45,472                                     | 8,058        | 56,072                     | 37,504                                     | 185,880                                       | 41,630                                   |
| 5,000.....                         | 7                                                    | 6,589              | 17,685                                     | 1,044        | 13,450                     | 4,913                                      | 47,822                                        | 11,048                                   |
| 10,000.....                        | 5                                                    | 9,951              | 15,301                                     | 1,936        | 18,255                     | 26,171                                     | 75,814                                        | 8,786                                    |
| 50,000.....                        |                                                      |                    |                                            |              |                            |                                            |                                               |                                          |
| 100,000.....                       |                                                      |                    |                                            |              |                            |                                            |                                               |                                          |
| Total..                            | 3,747                                                | 106,672            | 159,260                                    | 20,171       | 134,570                    | 127,365                                    | 580,964                                       | 128,271                                  |

Notes: These tables show examples of raw data from the SOI for the year 1945. Panel A is a screenshot of the tabulation by asset size bins for the aggregate economy. Panel B is a screenshot of the tabulation by asset size bins for one industry. Tabulations by asset size before 1960 did not include estimates for the small fraction of corporations that did not report balance sheets; see Internet Appendix [IA2.1](#) for detailed discussions and checks.

approximates a certain amount (e.g., top 1%). For instance, if the total number of businesses is  $N$  and the number of businesses in the top  $k$  bins adds up to less than  $0.01N$  (whereas the top  $k + 1$  bins add up to

more than  $0.01N$ ), then we take all the businesses in the top  $k$  bins and add  $(0.01N - \sum_{i=1}^k n_i)/n_{k+1}$  fraction from the  $k + 1$ th bin (where  $n_i$  denotes the number of businesses in the  $i$ th bin). In other words, we take all businesses in the top  $k$  bins and fill in the residual from the  $k + 1$ th bin. Overall, top shares using all methods are similar, as shown in Figure IA1. The raw correlation among the top 1% share estimated using generalized Pareto, generalized lognormal, and directly adding up bins is over 0.99. The benefit of interpolating distributions is that we do not have missing values for the small fraction of industry-years where the top bin has more than say 1% businesses; the benefit of adding up bins is that we can obtain other attributes of the top businesses (e.g., adding up their profits). We use generalized Pareto as the default, and use adding up bins when we need to measure other attributes of top businesses.

For the measure of concentration, we use the top 1% share as the baseline. The benchmark for evaluating the disparity of business size is the top  $x\%$  share in a number of models (Aghion et al., 2022; Hsieh and Rossi-Hansberg, 2022). Analyzing top percentiles is also the standard approach in research on household income and wealth inequality (Piketty and Saez, 2003; Saez and Zucman, 2016; Kuhn, Schularick, and Steins, 2020; Smith et al., 2019). Some models deliver the share of the top  $N$  businesses with a given  $N$  as a metric of concentration, but pinning down the right  $N$  across different settings can be challenging. For example, results using the same  $N$  may not be comparable across different levels of aggregation (e.g., top 500 firms is a small fraction for the aggregate economy but it can be a big fraction for a particular industry), or across industries that vary substantially in size. Table 2 shows the number of corporations in the aggregate and in the main sectors, and the *SOI Bulletin* publications provide extensive descriptions of the business population (Harris and Szeftlinski, 2007). Another concentration measure is the Herfindahl-Hirschman Index (HHI). This measure is too data intensive to be feasible for us because it requires data on the size of every single firm. To make sure the results of top 1% shares are not affected by small and extraneous firms coming in or out of the sample (therefore changing the total number of firms), we present the top 1% as a share of the top 10% as well, which should not be affected when the right tail is Pareto.<sup>6</sup> We also provide robustness checks using a fixed number of top businesses in Section 3.1.

For businesses with subsidiary affiliates, the SOI reports consolidated affiliates as one entity.<sup>7</sup> We follow IRS publications to refer to an entity in the SOI tabulations as a business or a corporation (see Petska and Wilson (1994), Harris and Szeftlinski (2007), and other *SOI Bulletin* publications). We explain consolidation rules in detail in Internet Appendix IA2 and provide a summary here. First, the consolidation threshold was 95% ownership of an affiliate before 1954 and 80% afterwards. The consolidated filing privilege is granted to all affiliated domestic corporations except regulated investment

<sup>6</sup>Specifically, if small and extraneous firms come in (out) of the data, the total number of firms in the top 1% will increase (decrease). Thus the top 1% share can increase (decrease), as the small firms have little impact on the total value of the denominator while the numerator will include more/less firms. To make sure our results are not affected by this issue, we can calculate the top  $x\%$  as a fraction of the top  $y\%$  (e.g., the top 1% as a share of the top 10%). One can show that for Pareto distributions, this relative share only depends on  $x/y$  and the tail coefficient  $k$ . In other words,  $\text{top } 1\% / \text{top } 10\% = \text{top } 0.01N / \text{top } 0.1N$  is invariant to the total number of firms  $N$ .

<sup>7</sup>For instance, the SOI in 2013 (as well as in other years) writes: "A consolidated return filed by the common parent company was treated as a unit and each statistical classification was determined on the basis of the combined data of the affiliated group."

Table 2 – Number of Corporations (000)

| Sector        | 1920 | 1930 | 1940 | 1950 | 1960  | 1970  | 1980  | 1990  | 2000  | 2010  |
|---------------|------|------|------|------|-------|-------|-------|-------|-------|-------|
| All           | 314  | 463  | 473  | 629  | 1,141 | 1,665 | 2,710 | 3,717 | 5,045 | 5,814 |
| Agriculture   | 9    | 10   | 8    | 8    | 17    | 37    | 81    | 126   | 141   | 137   |
| Construction  | 10   | 19   | 16   | 28   | 72    | 139   | 272   | 407   | 598   | 718   |
| Finance       | 79   | 137  | 143  | 172  | 334   | 406   | 493   | 609   | 748   | 892   |
| Manufacturing | 78   | 92   | 86   | 116  | 166   | 198   | 243   | 302   | 321   | 281   |
| Mining        | 18   | 12   | 10   | 9    | 13    | 14    | 26    | 40    | 33    | 40    |
| Services      | 17   | 38   | 41   | 55   | 121   | 281   | 671   | 1,029 | 1,791 | 2,273 |
| Trade         | 79   | 131  | 140  | 209  | 356   | 518   | 800   | 1,023 | 1,182 | 1,226 |
| Utilities     | 21   | 22   | 22   | 26   | 44    | 67    | 111   | 160   | 210   | 246   |

*Notes:* This table shows the number of corporations (in thousands) at the beginning of each decade, for the aggregate economy and the main sectors. The main sectors largely correspond to SIC codes 01-09 (agriculture), 10-14 (mining), 15-17 (construction), 20-39 (manufacturing), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

companies (RICs), real estate investment trusts (REITs), tax-exempt corporations, Interest Charge Domestic International Sales Corporations (IC-DISCs), and S-corporations. Second, consolidation was mandatory from 1918 to 1921 and voluntary after 1922, with the exception of 1934 to 1941 when consolidated filings were not allowed for most corporations. In recent decades at least, eligible firms generally elect to consolidate (Mills, Newberry, and Trautman, 2002), given more favorable treatments when consolidated (e.g., when consolidated the sales among affiliates do not generate taxes, and gains and losses across affiliates can be netted). Before 1964, there was often a small surtax on consolidated returns. In Internet Appendix IA2, we use SOI data to show the prevalence of consolidated filings over time, and examine the impact on our concentration estimates. We observe a decrease in the prevalence of consolidated filings between the early 1930s and the early 1940s, and then an increase between the 1960s and the 1980s (returning to the level observed in the early 1930s). Overall, the trend of rising concentration remains within each regime of consolidation filings.

For the reporting of financial information, firms provide their balance sheets (assets and liabilities) in Section L of Form 1120 and are instructed to use “the accounting method regularly used in keeping the corporation’s books and records” (see Form 1120 instructions). In other words, balance sheet items in Form 1120 (and correspondingly the SOI) largely follow what companies do for financial statements (with some possible differences such as the treatment of foreign affiliates and special purpose vehicles). Mills, Newberry, and Trautman (2002) provide detailed discussions about the relationship between financial information in SOI data and in firms’ annual reports. For assets and sales, we show in Section 3.2 that the size of top 500 firms estimated from our SOI data is similar to that calculated using Compustat data, so SOI data are in line with financial statement data. For net income, the SOI uses tax depreciation but the concentration series by net income is not our primary focus. Section 3.3 compares net income in SOI and NIPA (where the BEA makes adjustments to use economic depreciation instead), and we find the results are similar in the aggregate. Overall, reporting differences are unlikely to drive the main time trends we observe in Section 3 (given the high consistency among concentration trends

by assets, receipts, and net income).

In recent years, large asset managers (e.g., Vanguard, Blackrock) hold an increasing fraction of shares across multiple companies (see [Schmalz \(2018\)](#) and [Backus, Conlon, and Sinkinson \(2019\)](#) for reviews of the common ownership literature). Such common ownership driven by large index funds is unlikely to affect most of our sample period (before index funds became prominent) and sample firms (the vast majority are not publicly listed). Additionally, to our understanding the common ownership literature has not found strong evidence that different companies owned by the same investors directly coordinate their production activities. Since our focus is the organization of production activities, we treat a firm as a production enterprise in the usual way.

Like the national accounts, the SOI focuses on production activities in the U.S. (e.g., receipts in the SOI are similar to the convention of gross output in the national accounts). This is the natural realm for our analyses of concentration in production activities in the U.S. economy. For assets, the SOI data cover businesses incorporated in the U.S.; affiliates of foreign companies incorporated in the U.S. are included as U.S. corporations (they also do not count towards imports in the national accounts). For receipts, the SOI data cover U.S. corporations and foreign corporations with U.S. business activities (only income connected with conducting businesses in the U.S. is included); they cover exports but few imports. We discuss issues related to trade activities between U.S. and other countries in [Section 4.2](#). We also perform checks relating to foreign affiliates of U.S. businesses in [Section 3.2](#).

### 3 100 Years of Corporate Concentration

In this section, we present the evolution of top businesses' shares in the past century. We show the results for the aggregate economy and for different sectors in [Section 3.1](#). We provide additional checks of the concentration trends in [Section 3.2](#). We discuss long-run trends of other outcomes (e.g., profitability) in [Section 3.3](#).

#### 3.1 Basic Results

We begin by studying the population of corporations every year. We present results for the aggregate economy, the main sectors, and the subsectors.<sup>8</sup> We also present results including noncorporations when relevant data are available.

**Aggregate** [Figure 1](#) presents the trends for the aggregate economy. The green line with circles shows the share of top corporations by net income in total corporate net income (restricting to those with positive net income); the SOI tabulated corporations by net income in the early years, but stopped doing so after 1975. The red line with diamonds shows the share of top corporations by receipts in

---

<sup>8</sup>Analyses at the industry level mainly aim to group firms that share similar production activities (e.g., chemical manufacturing has different production processes from hotels). We do not stipulate that industry classifications map into product markets; as mentioned before, our focus is concentration in the production activities in the economy (not concentration in the market for a particular product or location).

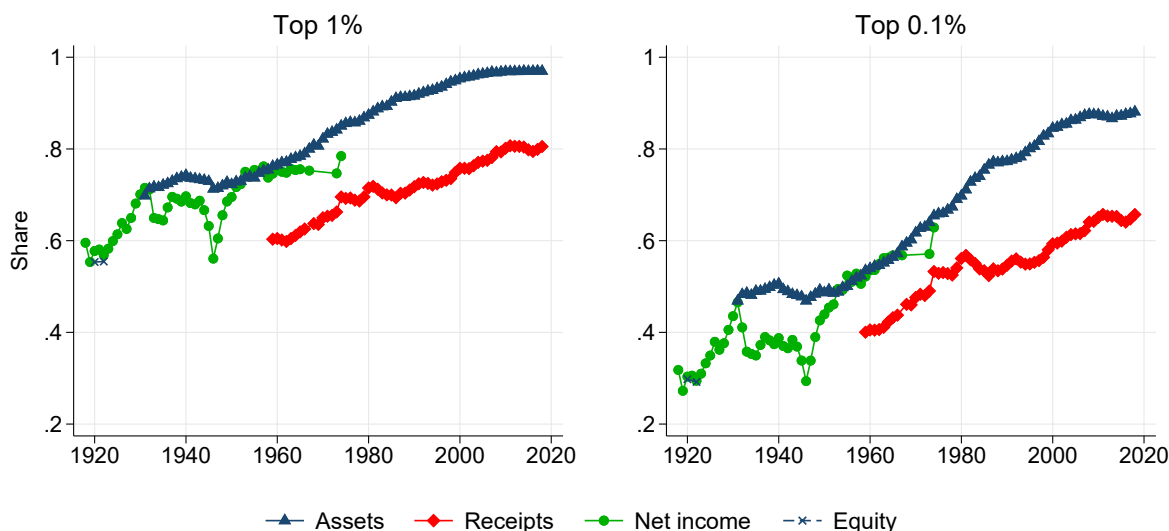


Figure 1. Top 1% and 0.1% Shares: All Corporations

*Notes:* This figure shows the shares of the top 1% (left panel) and the top 0.1% (right panel) corporations. The blue line with triangles shows the share of assets accounted for by top businesses sorted on assets. The red line with diamonds shows the share of receipts (sales) accounted for by top businesses sorted on receipts. The green line with circles shows the share of net income accounted for by top businesses sorted on net income (restricting to those with positive net income). The blue cross shows the share of equity capital accounted for by top businesses sorted on equity. See Internet Appendix [IA2.3](#) for details about variable construction.

total corporate receipts; the SOI started tabulating corporations by receipts in 1959. The blue line with triangles shows the share of top corporations by assets in total corporate assets; the SOI has consistently tabulated corporations by assets since 1931. In addition, for two years in the early 1920s the SOI tabulated corporations by their size of equity capital, and the share of equity accounted for by top corporations sorted on equity is shown using blue crosses; these data points align closely with the top share by net income that covers those early years. Overall, we observe consistent results across the three main types of size tabulations. Top shares by assets and by receipts have correlations over 0.9, and top shares by net income have correlations of around 0.7 with the other two series.

Figure 2 presents two more aggregate trends: the share of the top 1% among the top 10% (left panel) and the share of the top 0.1% among the top 1% (right panel). These additional series of relative top shares show the evolution of the far right tail of the size distribution. As discussed in Section 2, they also address the concern that some small firms at the bottom of the size distribution may not be very active, but they may affect the number of firms and correspondingly the top 1% share; the relative top shares should not be affected by this concern.

Interestingly, the secular trend of rising corporate concentration differs from the evolution of the top 1% and 0.1% household income and wealth shares in the U.S., which decreased between the 1920s and the 1970s and increased afterwards (Piketty and Saez, 2003; Saez and Zucman, 2016). In principle, whether corporate concentration and household inequality are linked depends on several factors. First, it depends on the extent to which the large businesses' revenues and profits disproportionately benefit a small number of individuals (e.g., due to concentrated equity ownership (Kuhn, Schularick, and



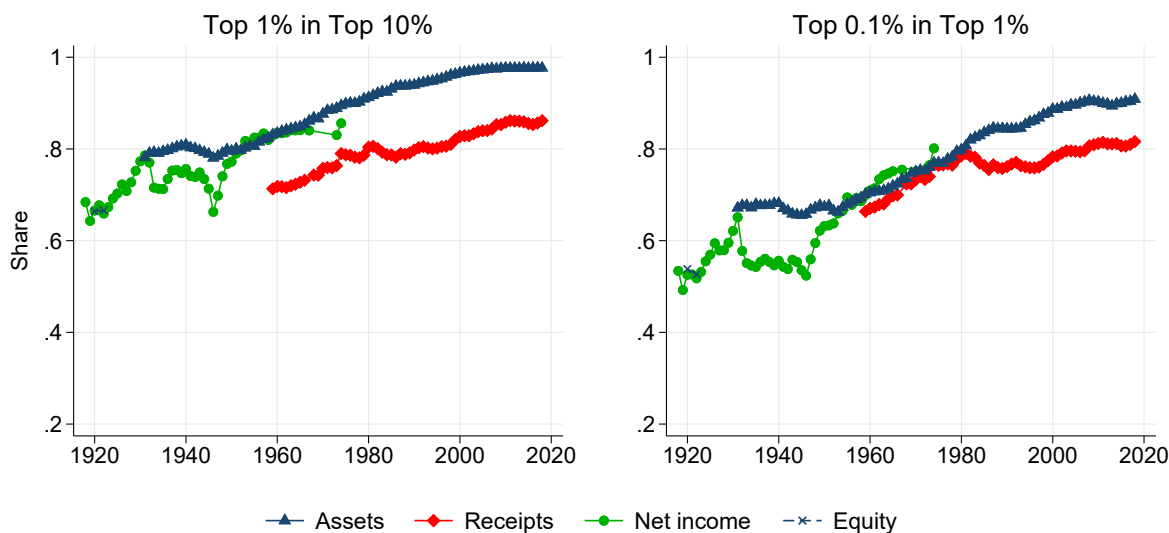


Figure 2. Relative Shares among Top Corporations

*Notes:* This figure shows the shares of the top 1% corporations among the top 10% corporations (left panel) and the top 0.1% corporations among the top 1% corporations (right panel). The blue line with triangles shows the share of assets accounted for by top businesses sorted on assets. The red line with diamonds shows the share of receipts accounted for by top businesses sorted on receipts. The green line with circles shows the share of net income accounted for by top businesses sorted on net income (restricting to those with positive net income). The blue cross shows the share of equity capital accounted for by top businesses sorted on equity. See Internet Appendix IA2.3 for details about variable construction.

Steins, 2020) or high executive compensation (Frydman and Saks, 2010)), rather than households more generally (e.g., if all households hold the market portfolio). Second, household inequality is also driven by redistribution policies (e.g., taxation), education, and many other forces.

**Main sectors and subsectors** We present results for the main sectors (around the one-digit SIC level) in Figure 3 and the subsectors (around the two-digit SIC level) in Figure 5. For the main sectors, we have size bins of corporations by assets, receipts, and net income. For the subsectors, tabulations are most comprehensive for size bins by assets, which also have the longest time series. Accordingly, we use the asset share of top corporations by assets as the main series in industry-level analyses.

Figure 3, Panel A, shows that concentration (as represented by the top 1% share) has been rising over the past century in most of the main sectors. The series by assets, receipts, and net income display consistent patterns. Figure 3, Panel B, focuses on concentration by assets, and shows that the results are similar for the share of the top 1% in the top 10%. Indeed, the share of the top 1% businesses in the top 10% is more than 0.98 correlated with the top 1% share, and all of our subsequent results about the top 1% hold for this series as well. Figure 3 also indicates that the timing for rising concentration varies across industries. The rise in concentration was stronger in earlier years for manufacturing and mining, and stronger in later years for services and trade (retail and wholesale). Panel A of Figure IA2 further visualizes the differences in timing. For each main sector, the solid blue circles show the change in the top 1% asset share between the 1930s and the 1970s, and the hollow red diamonds show the change between the 1970s and the 2010s. Table 3 provides a tabulation of the average top 1% asset share in



Table 3 – Top 1% Asset Shares: Average by Decade

|               | 1930s | 1940s | 1950s | 1960s | 1970s | 1980s | 1990s | 2000s | 2010s |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| All           | 0.72  | 0.73  | 0.74  | 0.79  | 0.85  | 0.90  | 0.93  | 0.96  | 0.97  |
| Agriculture   | 0.33  | 0.37  | 0.36  | 0.31  | 0.31  | 0.33  | 0.41  | 0.48  | 0.47  |
| Construction  | 0.42  | 0.33  | 0.34  | 0.37  | 0.43  | 0.50  | 0.49  | 0.58  | 0.63  |
| Finance       | 0.66  | 0.72  | 0.72  | 0.76  | 0.82  | 0.88  | 0.92  | 0.96  | 0.96  |
| Manufacturing | 0.67  | 0.65  | 0.70  | 0.77  | 0.85  | 0.89  | 0.91  | 0.94  | 0.95  |
| Mining        | 0.54  | 0.50  | 0.59  | 0.68  | 0.78  | 0.85  | 0.89  | 0.92  | 0.93  |
| Services      | 0.46  | 0.46  | 0.46  | 0.47  | 0.57  | 0.68  | 0.80  | 0.88  | 0.87  |
| Trade         | 0.49  | 0.47  | 0.44  | 0.46  | 0.54  | 0.66  | 0.74  | 0.80  | 0.86  |
| Utilities     | 0.82  | 0.81  | 0.87  | 0.92  | 0.95  | 0.96  | 0.96  | 0.97  | 0.97  |

*Notes:* This table shows the average top 1% asset share in each decade for the aggregate economy and the main sectors. The main sectors largely correspond to SIC codes 01-09 (agriculture), 10-14 (mining), 15-17 (construction), 20-39 (manufacturing), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

each decade for the aggregate and the main sectors.<sup>9</sup>

Figure 4 delineates the asset shares (Panel A) and receipt shares (Panel B) across the full distribution by size (top 0.1%, top 0.1% to 1%, top 1% to 10%, top 10% to 50%, and bottom 50%). First, this figure shows that most of the expansion of the top share is driven by the top 0.1%. Nonetheless, our data by size bins measure the top 1% shares more accurately; sometimes the top 0.1% has too few businesses, so we need to rely more on interpolation (e.g., size bins by receipts are less granular at the main sector level and Panel B shows that the interpolated top 0.1% shares by receipts have one-off jumps when the granularity of the top bin changes). Accordingly, we use the top 1% share as the main variable in our analysis, bearing in mind that much of its expansion can be driven by the top 0.1%. Second, Figure 4 also shows the expansion of the top businesses primarily reduces the shares of businesses in the middle of the distribution, and the bottom 50% is too small in value terms in any case.

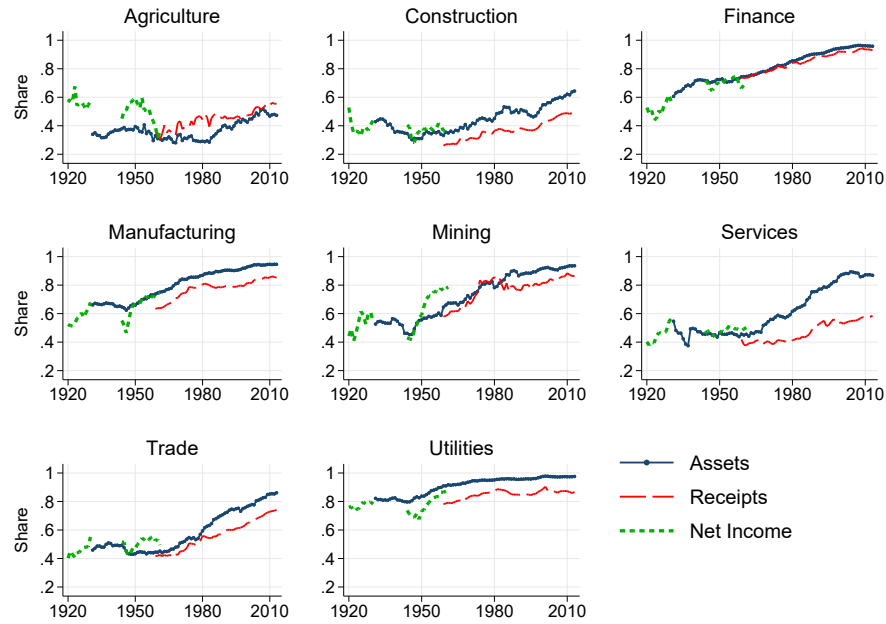
For the subsectors, Figure 5 shows the asset share of the top 1% businesses.<sup>10</sup> Similarly, the persistent rise in concentration is common in many industries, but the timing can differ. Panel B of Figure IA2 shows the change in the top 1% asset share between the 1930s and the 1970s and the change between the 1970s and the 2010s for each subsector. We investigate the timing in more detail in Section 4 to shed light on the mechanisms behind rising corporate concentration.<sup>11</sup>

<sup>9</sup>The level of top business shares can be higher in the aggregate than in most industries because some industries have more large firms and more concentrated industries may also have more large firms. For the *change* in aggregate top business shares, rising concentration in different industries in general, higher growth of industries with larger firms, and higher growth of industries with more concentration can all play a role.

<sup>10</sup>For all subsector analyses, we exclude “Finance: Holding Companies,” which includes RICs and REITs as these companies are the exceptions where consolidated filings are not allowed.

<sup>11</sup>While most industries experienced noticeable increases in concentration over time, the ranking in the cross section remains stable. For instance, the rank correlation between top 1% asset shares in the 1930s and those in the 2010s is over 0.9 among main sectors and around 0.7 among subsectors. This phenomenon suggests that industries differ persistently in the organization of production. The cross-industry dispersion of the top 1% asset share has decreased over time, as the top shares are bounded from the above for industries that were already concentrated in the early decades.

Panel A. Top 1% Shares: By Assets, Receipts, and Net Income



Panel B. Top 1% Shares: By Assets

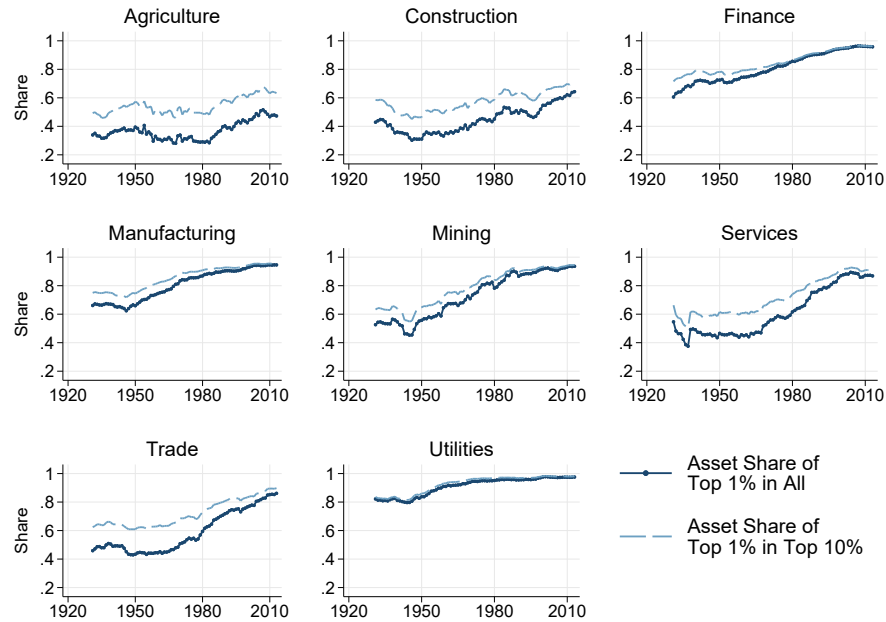
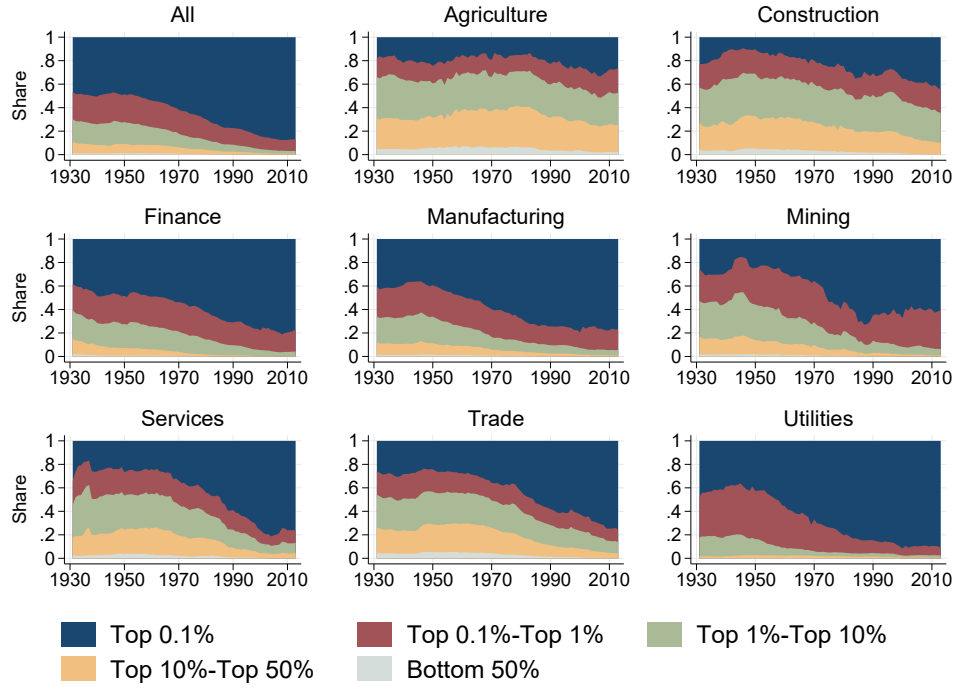


Figure 3. Top 1% Shares: Main Sectors

*Notes:* This figure shows the share of the top 1% corporations among the main sectors. In Panel A, the solid blue line shows the share of assets accounted for by top businesses sorted on assets. The dashed red line shows the share of receipts accounted for by top businesses sorted on receipts. The dotted green line shows the share of net income accounted for by top businesses sorted on net income (restricting to those with positive net income). In Panel B, the solid line repeats the share of the top 1% corporations by assets in total corporate assets, and the dashed line shows the share of the top 1% corporations by assets in the top 10% corporations by assets. The main sectors largely correspond to SIC codes 01-09 (agriculture), 10-14 (mining), 15-17 (construction), 20-39 (manufacturing), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

### Panel A. By Assets



### Panel B. By Receipts

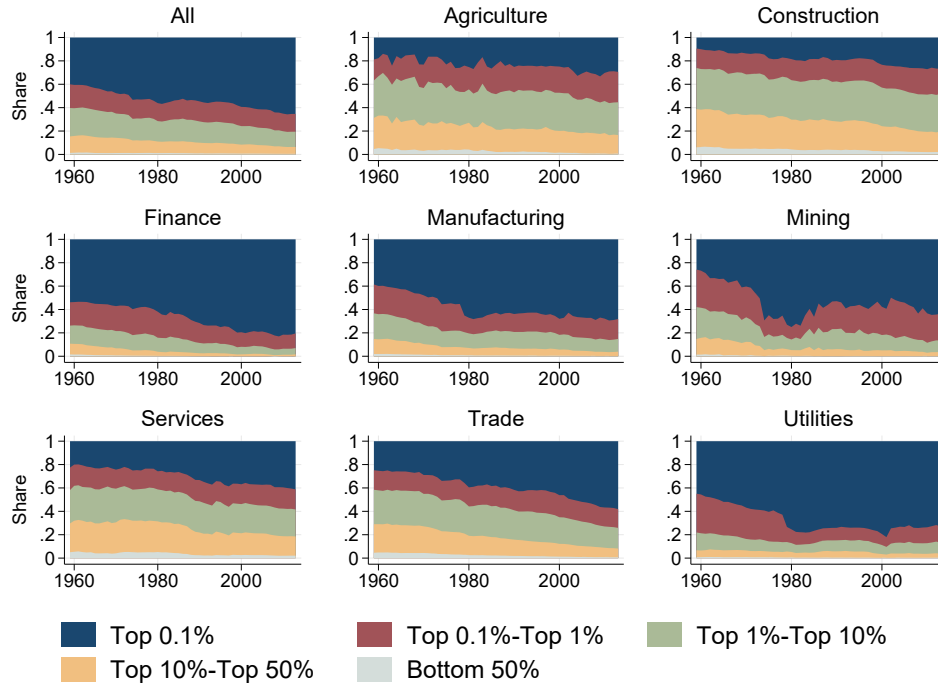


Figure 4. Full Distribution

Notes: Panel A shows the asset shares across the entire distribution of corporations by asset size: top 0.1% by assets (dark blue), top 0.1% to top 1% by assets (red), top 1% to top 10% by assets (green), top 10% to top 50% by assets (yellow), and bottom 50% by assets (gray). Panel B shows the receipt shares across the entire distribution of corporations by receipt size: top 0.1% by receipts (dark blue), top 0.1% to top 1% by receipts (red), top 1% to top 10% by receipts (green), top 10% to top 50% by receipts (yellow), and bottom 50% by receipts (gray).

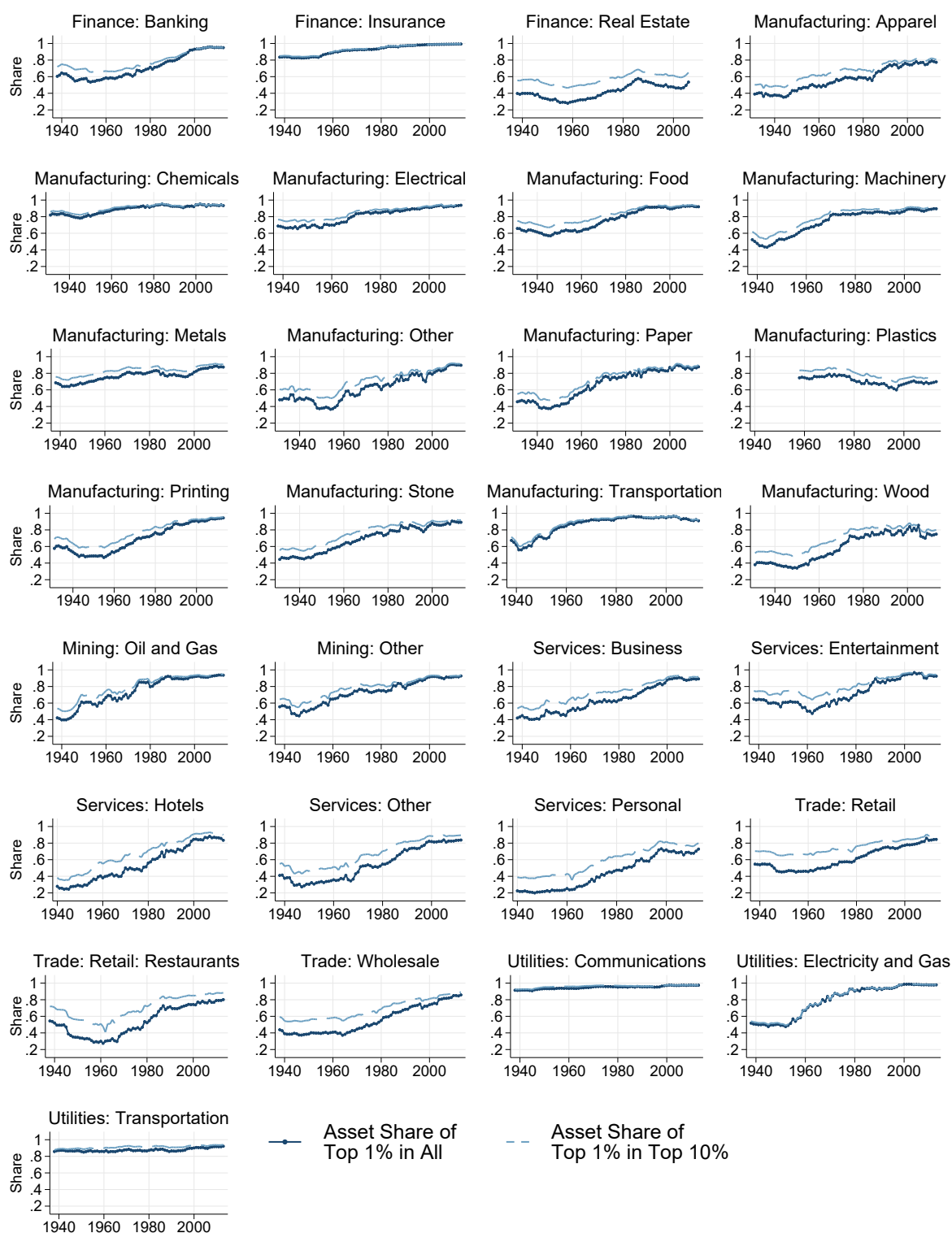


Figure 5. Top 1% Asset Shares: Subsectors

Notes: This figure shows the asset share of the top 1% corporations by assets in the subsectors. The solid line shows the share among all corporations and the dashed line shows the share among the top 10% corporations.

Finally, Figure IA3 shows examples of further breakdowns within our subsectors. Here we decompose mining into coal, metal, nonmetallic, and oil and gas in Panel A, and decompose retail into different types of retail in Panel B. The long-run trends of rising concentration also hold at this level of disaggregation. These further breakdowns are more granular than industries in datasets such as NIPA, which we rely on for subsequent analyses (about industry features that accompany rising concentration), so we do not include them in the subsectors in the main text. Further breakdowns of some other subsectors are not consistently available in the early years. For concentration in even narrower industries, the manufacturing census provides some historical information at the four-digit SIC level for census years, which we examine in Section 3.2. Overall, the long-run trends of rising concentration are present at various levels of aggregation.

**Including noncorporations** As explained in the Introduction, our objective is to study the concentration of production (i.e., the extent to which a small set of top businesses account for a large share of production activities). The SOI tabulations of businesses by size bins have comprehensive coverage of corporations (both C-corporations and S-corporations). In the following, we check that the concentration trends among corporations plus noncorporations (partnerships and proprietorships) are similar, using additional data on the size distribution of noncorporations when these data are available.

As an overview of the economic activities of corporations and noncorporations, Figure IA4 plots corporations' share in overall business receipts (by corporations, partnerships, and nonfarm proprietorships), using SOI data on the receipts of corporate and noncorporate sectors (see Internet Appendix IA2.1 for details). In the aggregate, the corporate share rose from 70% in the early decades to 90% in the 1980s, and decreased gradually to 80% since then; these trends are consistent with several studies showing that noncorporations have become more important since the 1980s (Clarke and Kopczuk, 2017; Kopczuk and Zwick, 2020). Among the main sectors, the corporate share is high for industries such as finance and manufacturing, as well as mining and utilities prior to the 1990s; it was lower in the earlier decades for construction, services, and trade. In our baseline results, rising concentration was stronger in manufacturing and mining before the 1980s, and the share of corporations was high and stable in these sectors during that period; rising concentration was stronger in services and retail/wholesale after the 1980s, and the share of corporations was also high and stable in these sectors during that period. In other words, at the industry level, rising concentration in our baseline results mainly occurred in time periods where corporations were dominant in output (so incorporation should not be driving the results in these time periods).

We perform checks for top business shares including noncorporations in Figure 6. In some years, we have tabulations of noncorporations by size bins based on receipts, and the data sources are listed in Table IA7. In these years, we can directly derive top businesses' receipt shares among all businesses (corporations plus noncorporations): we rank all businesses by size of receipts, and obtain the share accounted for by the top businesses among all businesses. Figure 6 shows that the top 1% receipt share among all businesses (purple diamonds) is similar to the top 1% receipt share among corporations (blue circles) in our baseline results. Overall, the patterns of rising concentration are similar when noncorporations are included.

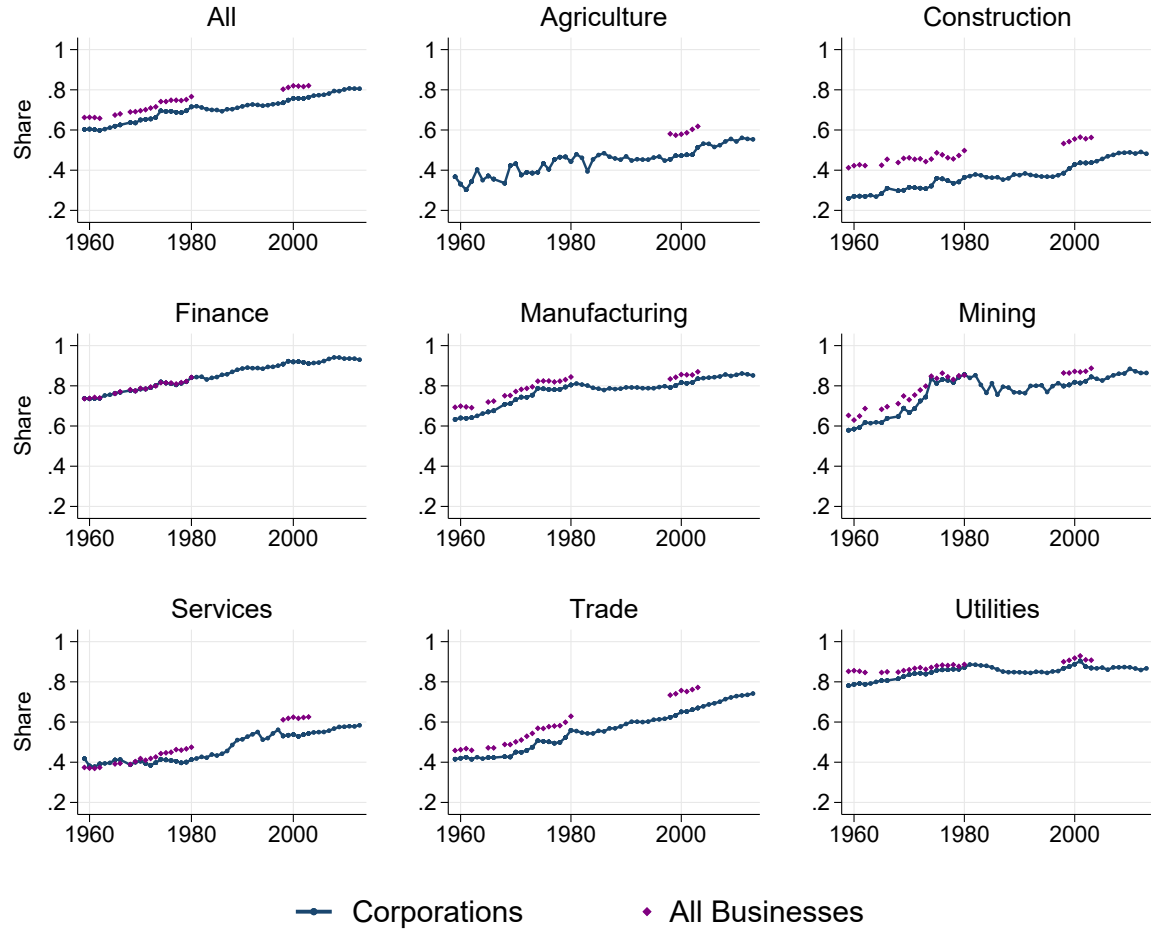


Figure 6. Top 1% Receipt Shares including Noncorporations

*Notes:* This figure shows robustness checks for top businesses' shares including noncorporations. The blue circles show the receipt share of the top 1% corporations by receipts among all corporations. The purple diamonds show the receipt share of the top 1% businesses by receipts among all businesses (corporations plus noncorporations), for years when size bins of noncorporations are available. The top receipt share among all businesses (purple diamonds) is not available for finance in recent decades due to data errors in the SOI publications. See Internet Appendix [IA2.3](#) for details about variable construction.

**Top  $N$  businesses with a given  $N$**  Another metric for concentration is the share of the top  $N$  businesses with a given  $N$ . As discussed in Section 2, it is not necessarily clear what  $N$  to choose for different levels of aggregation and different industries. We present the share of the top 5,000 corporations in the aggregate and the top 500 corporations in the main sectors as an example. Furthermore, the top  $N$  share with a fixed  $N$  can be more sensitive to the prevalence of corporations versus noncorporations. Intuitively, if we compute the top  $N$  share among corporations, the value of the denominator tends to expand (shrink) as corporations become more (less) common. In comparison, if we compute the top  $x\%$  share among corporations like in our baseline results, the numerator adapts and the prevalence of corporations does not make much difference as shown above in Figure 6.

To account for the influence of noncorporations on the top  $N$  share, we collect data on the assets, receipts, and net income of noncorporations. As mentioned before, data on noncorporations *by size* are



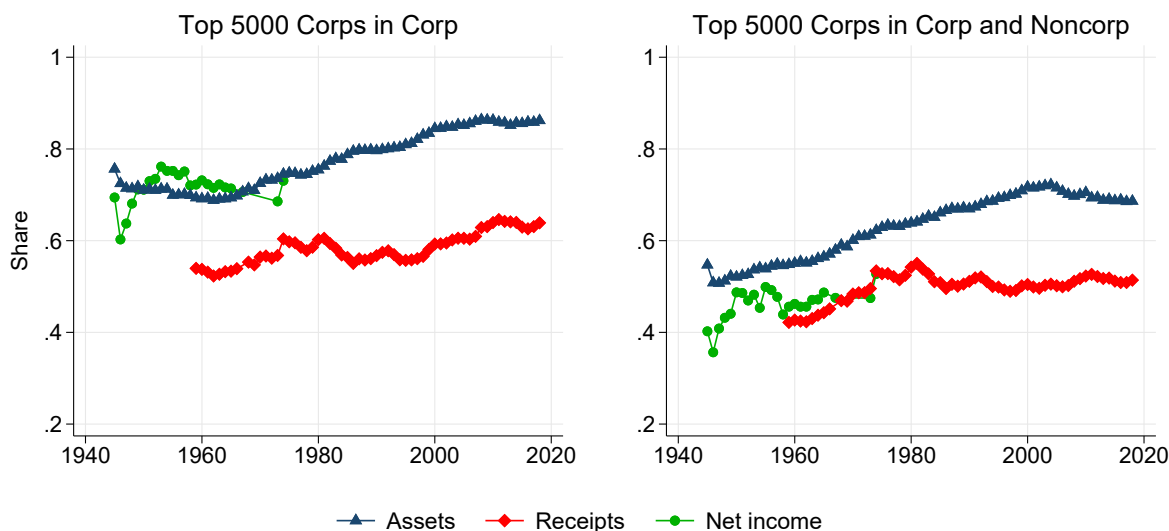


Figure 7. Shares of Top 5,000 Corporations

*Notes:* This figure shows the shares of the top 5,000 corporations among all corporations (left panel) and among all corporations plus noncorporations (right panel). The blue line with triangles shows the share of assets accounted for by top corporations sorted on assets. The red line with diamonds shows the share of receipts accounted for by top corporations sorted on receipts. The green line with circles shows the share of net income accounted for by top corporations sorted on net income (restricting to those with positive net income). See Internet Appendix [IA2.3](#) for details about variable construction.

sparse, so here we use top corporations for the numerator of the top share, and adjust the denominator with the total volume of noncorporations. For receipts and net income, we can obtain the total amount for partnerships and nonfarm proprietorships from SOI publications after 1945. For the aggregate top share by receipts (net income), we use the receipts (net income) of the top 5,000 corporations by receipts (net income) in the numerator, and include the receipts (net income) of corporations plus noncorporations in the denominator.<sup>12</sup> For assets, data on noncorporations are limited. We can make some estimates for noncorporations in the aggregate after 1945, using corporate and noncorporate assets from the Financial Accounts of the United States; we are not aware of data on noncorporations' assets at the industry level. For the aggregate top share by assets, we start with the share of the top 5,000 corporations by assets among total corporate assets in SOI data, and multiply it with the ratio of nonfinancial corporate assets to total nonfinancial business assets (corporate plus noncorporate) in the Financial Accounts (excluding farm assets and residential real estate to better match SOI data).<sup>13</sup>

For the aggregate economy, the left panel of Figure 7 shows the share of the top 5,000 corporations in all corporations, and the right panel shows the share of the top 5,000 corporations when we include both corporations and noncorporations in the denominator. The latter makes the rise of the top shares stronger before the 1970s (when the prevalence of corporations expanded), and weaker after the 1980s

<sup>12</sup>See Internet Appendix [IA2.1](#) for details about data sources. We focus on nonfarm proprietorships because the IRS stopped collecting data on farm proprietorships in 1981.

<sup>13</sup>Assets of the financial sector are not separated into corporate and noncorporate in the Financial Accounts. Therefore, our adjustment assumes that the ratio of corporate assets to noncorporate assets is similar in finance and the nonfinancial sector. This seems reasonable because the ratio of corporate receipts to noncorporate receipts (which we observe in SOI data) is similar in finance and the nonfinancial sector.

(when the prevalence of corporations declined). Meanwhile, after the 1980s large noncorporations have become more common, but they are not included in the numerator (due to limited data on noncorporations' size distribution), so the right panel may understate concentration in recent decades. For example, data on the size of noncorporations by receipts (analyzed in Figure 6 and only available in some years) suggest that the top 5,000 receipt share in the right panel of Figure 7 would be 10% (or 5 percentage points) higher around 2000 if the numerator includes the largest among all businesses; such an adjustment makes no difference before the 1980s. Overall, the long-run trends of rising concentration remains. Additionally, the number of potential entrepreneurs or the demand for goods and services may increase with the population, so we can also scale the number of businesses in the numerator with population. Figure IA5 shows the result where we have the top 5,000 corporations in the numerator at the beginning of the sample in 1945 and increase the number of top corporations in the numerator using cumulative population growth afterwards.

For the main sectors, we can only consistently obtain the receipts of noncorporations to adjust the denominator of the top  $N$  share. In Panel A of Figure IA6, we plot the share of the top 500 corporations by receipts among the receipts of corporations (solid line) and corporations plus noncorporations (dashed line). We also observe long-run increases of the top share, in manufacturing and mining before the 1980s and services, retail, and wholesale after the 1980s, consistent with the broad patterns in previous main sector level results. As before, including noncorporations in the denominator makes rising top share stronger when the corporate sector expands (earlier decades) and weaker when the corporate sector shrinks (later decades). For industries where large noncorporations have become more common among top firms after the 1980s (e.g., mining and utilities), the share of the top 500 corporations among corporations plus noncorporations may understate concentration in recent decades. For example, data on noncorporations by receipt size (analyzed in Figure 6 and only available in some years) suggest that the top 500 receipt share among corporations plus noncorporations around 2000 is similar to that around 1980 for mining and utilities if the numerator includes the largest among all businesses. In Panel B, we start with the top 500 corporations at the beginning of the sample and increase the number of top corporations in the numerator using cumulative population growth afterwards.

## 3.2 Additional Results

**Comparison with census data** We cross-check our data with census data, which report sales shares of the top 4, 8, 20, and 50 firms by sales, for manufacturing industries in census years since 1947 and other industries after 1982.<sup>14</sup> First, for the degree of concentration over time, we rely on the longer time series of census concentration ratios in manufacturing industries at the four-digit SIC level (analyzed in Pryor (2001), Peltzman (2014), Lamoreaux (2019) and Keil (2017) among others). In this dataset, we can take the average of census concentration ratios across these detailed industries. Figure IA7, Panel

---

<sup>14</sup>As explained in census documentations (e.g., "Concentration ratios in manufacturing industry, 1958"), the concentration ratios for each industry are based on the sales of establishments in that industry. Therefore, if a firm spans multiple industries, its establishments in different industries are treated as different observations: the sales of establishments in industry A (B) count towards the concentration ratio in industry A (B). In other words, a given firm can be split into several observations. Meanwhile, as explained in Section 2, the SOI data always treat a business as one observation, which belongs to the industry that represents the largest percentage of its total receipts.

A, displays the value-weighted average (solid line) and the equal-weighted average (dashed line). We observe a persistent increase in the sample period, especially for the value-weighted average.<sup>15</sup>

Second, for the degree of concentration across industries, we compare census concentration ratios in 2012 (with both manufacturing and non-manufacturing industries) with our estimates using SOI data. The census data in 2012 are available for two-digit to six-digit NAICS industries. The most granular SOI industries in 2012 map into roughly four-digit NAICS codes. Here we can estimate the sales shares of the top businesses by assets (by adding up the sales of top corporations by assets); we cannot rank businesses by sales in these granular industries where the SOI data only report size bins by assets. Accordingly, our measure is not exactly the same as the census measure. Figure IA7, Panel B, shows the cross-sectional relationship between the sales shares of top 20 businesses by sales in census data ( $x$ -axis) and the sales shares of top 20 businesses by assets in SOI data ( $y$ -axis). We observe a high degree of consistency in the cross section, with a correlation of 0.84. The level is also similar, with a mean (median) difference of about 0.03.

**Comparison with Compustat** We also present a comparison with aggregate top shares using Compustat data. In general, it is difficult to perform comprehensive analyses of business concentration using stock market data alone. First, stock market data miss large private firms while at the same time have some smaller growth firms. Second, the coverage of stock market data is very limited at the industry level. For example, in 1960, a quarter of our subsectors had less than 10 public companies and nearly one half had less than 20 companies. This makes it challenging to properly calculate top shares at the industry level using public firms (and the result can be sensitive to the listing decision of one particular firm). Several studies also show that industry-level concentration measures using Compustat data versus census data differ substantially (Keil, 2017; Ali, Klasa, and Yeung, 2008). Accordingly, we perform an aggregate comparison, and our objective is to check that the level of aggregate top shares is comparable in SOI and Compustat data. This analysis verifies that businesses in the SOI data are similar to those in financial statements (e.g., consolidation is performed properly in the SOI), and the SOI data reliably capture firms at the very top.

In Panel A of Figure IA8, we calculate the total assets of the top 500 firms by assets in Compustat, and compute their share in total corporate assets (from SOI data). We compare this series with the imputed share of the top 500 corporations by assets in total corporate assets using SOI data (the imputation can be imperfect in the later years when the top bin contained much more than 500 businesses). In Panel B, we calculate the total sales of the top 500 firms by sales in Compustat, and compute their share in total corporate receipts (from SOI data). We compare this series with the imputed share of the top 500 corporations by receipts in total corporate receipts using SOI data. Total assets and sales in Compustat generally include global activities; after around 1998, activities of foreign subsidiaries can be separated using Compustat data on geographical segments, but the segment data are less reliable (e.g., a segment can include North America as a whole) and unavailable for many companies. Accordingly, the selection issue of public firms will lead to a downward bias in top business shares using Compustat data, while

---

<sup>15</sup>Peltzman (2014) tabulates the equal-weighted average of the change in CR4 between 1963 and 1982, which is close to zero. This is consistent with the milder increase in the equal-weighted averages in Figure IA7, Panel A. In addition, the rise in concentration in this period is stronger among a broader set of firms (e.g., CR20 compared to CR4).

reporting global activities may lead to an upward bias.

Figure IA8 shows that the level of aggregate top 500 shares is similar using SOI data (solid line) and Compustat data (dashed line). Both display an upward trend since Compustat data became comprehensive in the 1960s, though the trend in Compustat data needs to be viewed with caution given the data coverage issues discussed above. Overall, the results show that our SOI data are reliable for capturing firms at the very top and a business in the SOI data should be similar to a firm in Compustat.

**Including international assets of U.S. companies** While we focus on production activities of businesses in the U.S., in recent years some U.S. firms may have shifted production assets to foreign subsidiaries (Auerbach, 2021). We perform checks that include the assets of U.S. businesses' foreign affiliates, using Activities of U.S. Multinational Enterprises compiled by the BEA. Since the early 1980s, the BEA data record the assets of foreign affiliates and information about their U.S. parents. According to the number of U.S. parents with foreign affiliates reported in these data, less than 1% businesses are multinational in all main sectors.

We perform checks including international assets under two assumptions. A stronger assumption is that all international assets belong to the top 1% businesses. This assumption seems reasonable as the average assets for U.S. parents with foreign affiliates are almost always larger than the average assets of the top 1% businesses in every main sector. A weaker assumption is that the top 1% businesses' share of international assets is the same as their share of domestic assets. It is unlikely that businesses outside of the top 1% account for a larger proportion of the international assets. In Figure IA9, the solid blue line shows the original top 1% asset share; the dashed red line shows the adjusted top 1% asset share where we allocate all international assets to the top 1% businesses; the dash-dotted purple line shows the adjusted top 1% share where we allocate international assets to the top 1% businesses and the rest according to their shares in domestic assets. The concentration trends including international assets are similar to our baseline results (international assets are less than 20% of domestic assets in most industries except manufacturing and services after the 2000s, and the ratio of international assets to the top 1% businesses' domestic assets has remained stable).

**Employment concentration** The SOI provides information on business size by financial metrics such as assets, receipts, and net income; it does not provide information on employment. Does employment concentration among U.S. firms also increase over time? Since 1979, the census database on Business Dynamics Statistics (BDS) tabulates the number of firms and their employment by employment bins. We can therefore estimate the employment shares of top firms by employment size in BDS data. We are not aware of earlier data on firm size distribution by employment.

Figure IA10 plots the aggregate share of the top 1% firms by employment in total employment using BDS data. First, the level of employment concentration is lower relative to the level of concentration measured by financial outcomes. Second, employment concentration displays a slight increase in the sample period (e.g., the top 1% share rose from 55% in 1979 to around 60% in the 2010s); the magnitude is modest and can be less visible as shown in Luttmer (2010). Autor et al. (2020) perform detailed analyses of employment concentration using census micro data since the 1980s for a wide range of industries. They show that industry-level employment concentration is lower than sales concentration,

and it increased in this period but less than sales concentration did. As they explain, top firms produce more with fewer workers and exhibit “scale without mass”; accordingly, employment concentration has a lower level and rises less as large firms expand. [Hubmer and Restrepo \(2022\)](#) and [Firooz, Liu, and Wang \(2022\)](#) also suggest that large firms have become more capital intensive and less labor intensive.<sup>16</sup>

### 3.3 Other Outcomes

Finally, we present several additional outcomes to provide further context of corporate activities during our sample period.

**Profitability** The SOI data provide a variety of financial information for corporations in each size bin, as shown by the examples in Table 1. Using this information, Figure 8 shows the profitability ratio (i.e., net income before tax over sales) for the top 1% businesses by assets (solid line) and the rest (dashed line). Several patterns emerge from this figure. First, the profitability ratio has fluctuated substantially over time; it does not exhibit a persistent long-run trend. Profitability in almost all sectors was low during the Great Depression; it then rebounded sharply in the 1940s, declined until the 1980s, and increased slightly afterwards. These trends are in line with the analyses of corporate profits since 1945 by [Barkai and Benzell \(2018\)](#). Second, profitability is higher among the top 1% businesses than among the remaining businesses, but the difference between these two groups does not display noticeable changes over time.

Because net income is affected by depreciation and tax rules for depreciation have changed over time, we also cross-check profitability in the SOI with that in the national accounts. The BEA begins with data from the SOI and then makes capital consumption adjustments so that corporate profits use economic depreciation (estimated by the BEA). Figure IA11 shows corporate profits according to SOI and BEA, both normalized by total receipts from SOI. The result shows that aggregate corporate profits from these two sources are similar.

Overall, the data show that corporate profitability has fluctuated over the past 100 years; it has not followed the same persistent trend as corporate concentration. Estimating markups is more challenging, as shown by the ongoing discussions in the literature ([Hall, 2018](#); [Traina, 2018](#); [De Loecker, Eeckhout, and Unger, 2020](#); [Foster, Haltiwanger, and Tuttle, 2022](#)). Nonetheless, existing estimates using different methods also do not find that markups increased before the 1980s.

**Asset ownership** We also examine whether changes in the balance sheet characteristics of larger and smaller businesses affect the rising concentration trends we observe (we focus on nonfinancial industries here since the balance sheet structure of financial services is substantially influenced by regulations). One possible concern is that maybe smaller firms lease more assets over time, so their book assets (which do not include most leases) will shrink relative to those of larger firms; this concern, however, should not affect concentration by sales. In Figure IA12, we plot the ratio of fixed assets on firms’ balance sheets relative to their total assets, since leasing mainly applies to fixed assets. We do not

---

<sup>16</sup>Similarly, [Lenin \(1916\)](#) wrote: “Concentration of production, however, is much more intense than the concentration of workers, since labour in the large enterprises is much more productive.”

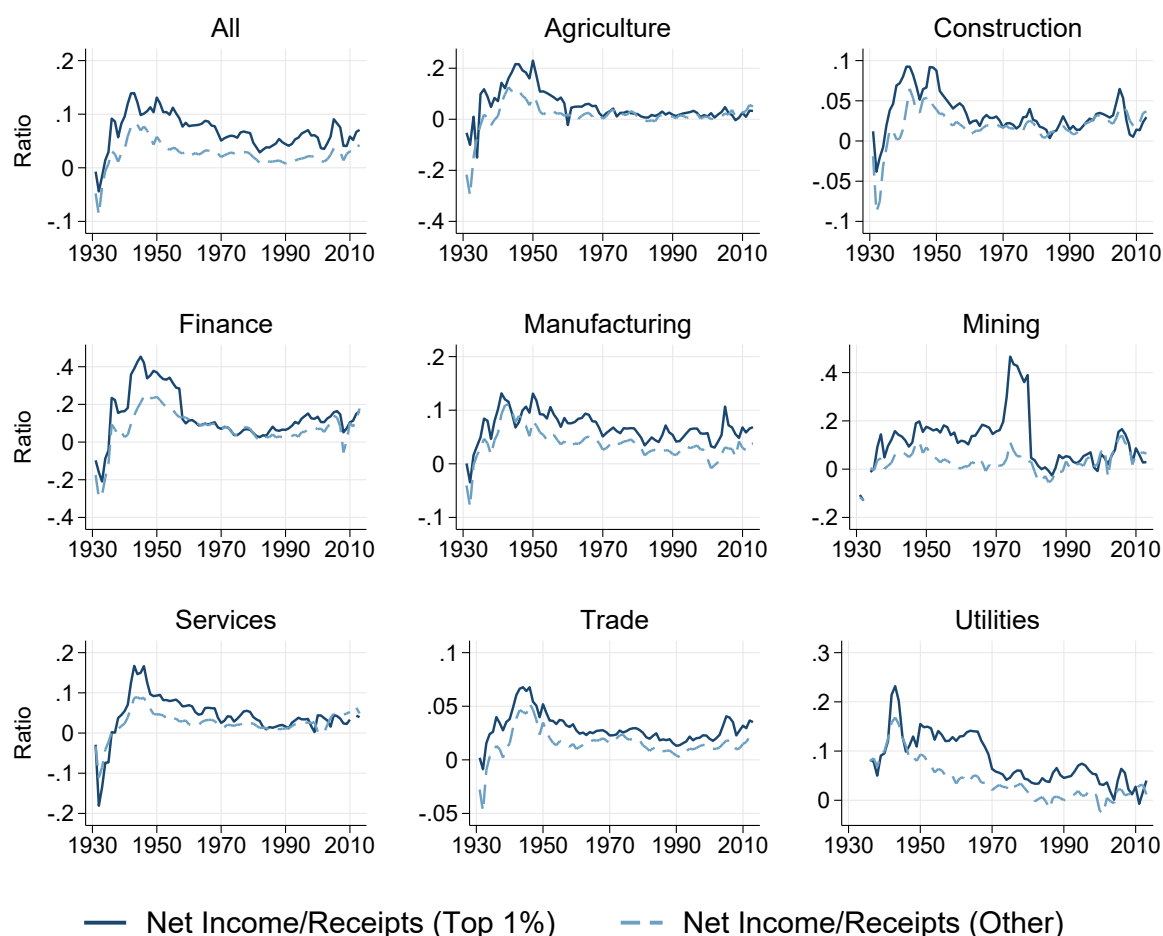


Figure 8. Profitability

*Notes:* This figure shows the profitability ratio (net income before tax over total receipts) for the top 1% corporations by assets (solid line) and the rest (dashed line). Here we need to use the adding up bins method discussed in Section 2 to obtain net income and receipts for these two groups (i.e., we add up the net income and receipts for each of these two groups). See Internet Appendix IA2.3 for details about variable construction.

observe different long-run trends for the ownership of fixed assets among the top 1% corporations by assets and the rest.

**Investment rate** Recent work postulates that the decline in corporate investment rates in the past few decades is linked to rising concentration (Gutiérrez and Philippon, 2017). In Figure IA13, we plot the long-run relationship between the investment rate (investment spending over asset stock using BEA fixed asset tables) and corporate concentration (top 1% asset share). We include investment rates using fixed assets alone (dashed line) and fixed assets plus intellectual property (dash-dotted line); even though annual investment spending in the BEA data started in 1901, the fixed asset stock only began in 1947. The investment rate in fixed assets shows a decline in many sectors, but the decline is less evident when intellectual property is included, in line with findings in Crouzet and Eberly (2021). Overall, Figure IA13 suggests that, over the long run, there does not appear to be a strong association between changes in investment rates and concentration.



**Labor share** Several studies use census data across different industries to document that falling labor shares since the 1980s are associated with concurrent increases in concentration (Autor et al., 2020; Barkai, 2020; Ganapati, 2021). Over the long run, the labor share in most industries did not decline before the 1980s (Elsby, Hobijn, and Şahin, 2013), even though we observe a secular increase in corporate concentration. As Hubmer (2022) points out, the long-run evolution of the labor share could be affected by other forces such as preferences, which are beyond the focus of our study. Accordingly, at the moment we do not dive into the long-run relationship between the labor share and concentration.

**Entry rate** Recent work also uses census data to document declining firm entry rates since the 1980s (Decker et al., 2014a,b). We use census BDS data (available since 1978) for firm entry rates (i.e., the share of new firms) across industries, and examine the relationship with concentration trends. Figure IA14 shows that rising concentration is generally correlated with decreasing entry rates. However, this relationship can be consistent with multiple mechanisms. For instance, stronger economies of scale can increase concentration and reduce entry. Changes in regulatory policies may also increase concentration and reduce entry. Accordingly, entry rates per se may not provide enough information for the underlying mechanisms.

**Mergers** We do not focus on analyzing mergers for three reasons. First, the prevalence of mergers can be affected by a number of mechanisms, including economies of scale, antitrust, and other regulations. Second, Holmstrom and Kaplan (2001) collect data on aggregate merger volume as a percentage of GDP from 1968 to 1999 and do not find a strong trend over these decades. Third, it is difficult to obtain comprehensive long-run merger data at the industry level.<sup>17</sup>

## 4 Mechanisms

In this section, we discuss the leading hypotheses about rising concentration. These hypotheses have attracted considerable attention in recent research. We investigate the extent to which they align with the evolution of production concentration over the long run. It is inevitably challenging to pin down the exact cause of rising concentration, but certain mechanisms appear more consistent with the overall long-run trends and the timing of rising concentration in different industries.

### 4.1 Economies of Scale

A longstanding observation suggests that stronger economies of scale will increase concentration in various economic domains (Demsetz, 1973; Rosen, 1981; Frank and Cook, 1996; Kaplan and Rauh, 2013). For firms, following the Second Industrial Revolution, scalable production became more common, leading to the emergence of large industrial enterprises (Chandler, 1994). Over time, economies of scale spread among other industries as well (e.g., retail and services), especially with the advancement of modern IT (Brynjolfsson et al., 2008; Hsieh and Rossi-Hansberg, 2022; Aghion et al., 2022; Lashkari,

---

<sup>17</sup>For large mergers studied in merger waves such as the conglomerate boom, many represent firms in the top 0.1% buying firms in the top 1%, so the impact of such a merger on the top 1% share may not be substantial.



Bauer, and Boussard, 2022).<sup>18</sup> We present three sets of results that align with this mechanism. We then present a simple model to illustrate that new technologies enhancing the scalability of production can lead to the empirical results that we observe.

First, we examine the relationship between production concentration in an industry and its technological intensity. We begin with a general measure of technological intensity using the spending on R&D and IT (computer equipment and software). This measure follows the common view that R&D and IT enhance the scalability of production (Brynjolfsson et al., 2008; Haskel and Westlake, 2017; Crouzet and Eberly, 2019; Lashkari, Bauer, and Boussard, 2022).<sup>19</sup> We can obtain this general measure annually for all industries using BEA's fixed asset tables (available since 1901). We use top business shares among the subsectors, which largely map into industries in BEA datasets as shown in Internet Appendix IA2.2. Table 4 presents the results, for all industries in columns (1) to (4), and nonfinancial industries in columns (5) to (8) since finance subsectors have been more affected by regulation. We present regressions both in levels and using changes within an industry over the medium term. We also absorb common time trends in the even columns. For instance, column (4) and (8) zoom into the timing alignment: at a given point in time, those industries that experience more increases in concentration are also the ones that experience more increases in technological intensity.<sup>20</sup> Figure IA15 visualizes the relationship. We normalize R&D and IT spending with total investment in BEA fixed asset tables for the baseline analysis. Results are similar if we instead normalize with business receipts in each industry from SOI data, shown in Table IA1 (we use this variable as a robustness check since normalizing across different datasets may introduce more noise). Bessen (2020) analyzes census concentration ratios from 1997 to 2012, and finds a positive relationship with IT intensity as well.

We then perform additional analyses using long-run data on breakthrough patents constructed by Kelly et al. (2021) to measure technological innovations in production activities, which are available for manufacturing and mining subsectors plus agriculture, construction, and utilities. A patent represents a breakthrough if it is distinct from patents that came before, but followed by subsequent patents that are similar. We normalize the number of breakthrough patents and total patents by population as in the original study; results are similar if we use alternative normalization such as real GDP. Table IA3

---

<sup>18</sup>The measurement of economies of scale also has a long intellectual history. Stigler (1958) offers an in-depth discussion in an article titled "The Economies of Scale," which builds on comments by Milton Friedman at a 1955 NBER conference on "Business Concentration and Price Policy." Stigler (1958) suggests that examining the evolution of the business size distribution is perhaps the most straightforward method to reflect the extent of economies of scale. In particular, he suggests to "classify the firms in an industry by size, and calculate the share of industry output coming from each class over time. If the share of a given class falls, it is relatively inefficient, and in general is more inefficient the more rapidly the share falls."

<sup>19</sup>This general measure can encompass two possibilities. First, IT and R&D can be directly involved in technological changes that enhance economies of scale. For instance, industrial R&D played an important role in the development and commercialization of chemical products, and IT played an important role in the rise of large supermarket chains. Second, it is also possible that other forces increase the benefits of scale production (e.g., lower transportation costs, shifts in consumer preferences), and scale production entails more IT and R&D. For example, restaurant chains formed partly because eating out became more common and customers looked to save time during work days and travels. Managing such chains requires more IT and at times also more R&D (e.g., to ensure quality stability).

<sup>20</sup>In Table IA2, we address the concern that changes in the number of corporations may affect the top 1% share. We repeat the regressions in Table 4 controlling for the log change in the number of corporations. Our main results are similar; the coefficient on changes in the number of corporations is small and insignificant, which further verifies that these changes are not a main driver of rising concentration in our data.

Table 4 – Rising Concentration and Technological Intensity

|                                       | Top 1% Asset Share  |                     |                          |                     |                     |                     |                           |                     |
|---------------------------------------|---------------------|---------------------|--------------------------|---------------------|---------------------|---------------------|---------------------------|---------------------|
|                                       | Level               |                     | Change ( $\Delta_{20}$ ) |                     | Level               |                     | Change ( $\Delta_{20}$ A) |                     |
|                                       | (1)                 | (2)                 | (3)                      | (4)                 | (5)                 | (6)                 | (7)                       | (8)                 |
| Intensity of IT and R&D               | 0.579***<br>(0.083) | 0.370***<br>(0.079) |                          |                     | 0.552***<br>(0.092) | 0.338***<br>(0.083) |                           |                     |
| $\Delta_{20}$ Intensity of IT and R&D |                     |                     | 0.122**<br>(0.060)       | 0.082***<br>(0.028) |                     |                     | 0.190***<br>(0.061)       | 0.157***<br>(0.036) |
| Industries                            | All                 |                     |                          |                     | Nonfinancial        |                     |                           |                     |
| Year Fixed Effect                     | No                  | Yes                 | No                       | Yes                 | No                  | Yes                 | No                        | Yes                 |
| Obs                                   | 2,394               | 2,394               | 1,774                    | 1,774               | 2,173               | 2,173               | 1,613                     | 1,613               |
| R <sup>2</sup>                        | 0.30                | 0.15                | 0.02                     | 0.01                | 0.28                | 0.14                | 0.03                      | 0.02                |

*Notes:* This table shows industry-level regressions of the top 1% asset share on the investment intensity of IT and R&D. The intensity of IT and R&D is measured using the annual spending on IT (computer equipment and software) and R&D in BEA fixed asset tables, normalized by total annual investment spending (on fixed assets and intellectual property) in BEA fixed asset tables. For both left hand side and right hand side variables, we use their levels in columns (1), (2), (5), and (6), and their changes over twenty years in columns (3), (4), (7), and (8). Year fixed effects are included in the even columns. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in BEA fixed asset tables follows Table IA8. Standard errors are Driscoll and Kraay (1998) with twenty lags.  $R^2$  does not include fixed effects.

shows that higher intensity of breakthrough patents is correlated with more increases in concentration. The results are stronger in the subsample before the 1980s, which is the time period when rising concentration among manufacturing-related industries was strongest. The patent data also help address a reverse causality concern about the results above: maybe large businesses report more comprehensively their spending on R&D and IT, and shocks that happen to benefit large firms (even randomness) increase concentration as well as the industry-level investment intensity of R&D and IT. We control for the number of patents in the even columns in case large firms are more likely to register patents. Indeed, we find that breakthrough patents show much stronger results than the number of patents. Autor et al. (2020) examine the relationship between rising concentration and increases in the amount of patents in detailed manufacturing industries from 1982 to 2012, and find a positive relationship as well.

Second, we examine the relationship between production concentration and the prevalence of fixed operating costs. Several studies have been interested in detecting fixed operating costs using firms' financial statements (Anderson, Banker, and Janakiraman, 2003; De Loecker, Eeckhout, and Unger, 2020; Traina, 2018), especially through the decomposition of operating costs into costs of goods sold (COGS, often thought to be more variable) and selling, general, and administrative expenses (SG&A, often thought to be more fixed). We do not stipulate whether these categories represent entirely variable or fixed costs. Rather, we allow a portion of each category to be variable, and estimate this portion using the regression coefficient of annual log changes in COGS (SG&A) on log changes in sales among Compustat firms in each industry  $k$ , which we denote by  $\alpha_k$  ( $\beta_k$ ). For the average industry, we find that COGS (SG&A) changes by 0.75% (0.46%) for a 1% change in sales. We then estimate the proportion of variable costs for a Compustat firm  $i$  in a given year  $t$  as  $vc_{it}^k = (\alpha_k \text{COGS}_{it} + \beta_k \text{SG\&A}_{it}) / (\text{COGS}_{it} + \text{SG\&A}_{it})$ , and the proportion of fixed costs as  $fc_{it}^k = 1 - vc_{it}^k$ . Finally, we take the median value of  $fc_{it}^k$  for each

Table 5 – Rising Concentration and Fixed Cost Intensity

|                                 | Top 1% Asset Share  |                     |                          |                     |                     |                     |                          |                     |
|---------------------------------|---------------------|---------------------|--------------------------|---------------------|---------------------|---------------------|--------------------------|---------------------|
|                                 | Level               |                     | Change ( $\Delta_{20}$ ) |                     | Level               |                     | Change ( $\Delta_{20}$ ) |                     |
|                                 | (1)                 | (2)                 | (3)                      | (4)                 | (5)                 | (6)                 | (7)                      | (8)                 |
| Fixed Costs Share               | 0.349***<br>(0.075) | 0.293***<br>(0.063) |                          |                     | 0.366***<br>(0.067) | 0.313***<br>(0.057) |                          |                     |
| $\Delta_{20}$ Fixed Costs Share |                     |                     | 0.564***<br>(0.061)      | 0.399***<br>(0.056) |                     |                     | 0.658***<br>(0.060)      | 0.459***<br>(0.081) |
| Industries                      | All                 |                     |                          |                     | Nonfinancial        |                     |                          |                     |
| Year Fixed Effect               | No                  | Yes                 | No                       | Yes                 | No                  | Yes                 | No                       | Yes                 |
| Obs                             | 1,869               | 1,869               | 1,249                    | 1,249               | 1,699               | 1,699               | 1,139                    | 1,139               |
| R <sup>2</sup>                  | 0.07                | 0.07                | 0.03                     | 0.02                | 0.09                | 0.09                | 0.04                     | 0.02                |

Notes: This table shows industry-level regressions of the top 1% asset share on the intensity of fixed operating costs. We estimate the amount of fixed operating costs in total operating costs for Compustat firms, and take the median in each industry for each year. For both left hand side and right hand side variables, we use their levels in columns (1), (2), (5), and (6), and their changes over twenty years in columns (3), (4), (7), and (8). Year fixed effects are included in the even columns. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in Compustat follows Table IA6. Standard errors are Driscoll and Kraay (1998) with twenty lags.  $R^2$  does not include fixed effects. See Internet Appendix IA2.3 for details about variable construction.

industry  $k$  in a given year  $t$  among Compustat firms. Compustat firms are generally large and likely to utilize scalable technologies, so they can be useful for reflecting the relevance of scalable technologies with higher fixed costs. Compustat data on cost structure are widely available after the 1950s.<sup>21</sup> Table 5 shows that the top 1% share is positively correlated with the intensity of fixed operating costs. In this analysis, it is unlikely that the fixed cost share will mechanically increase if certain firms get larger. By construction, the fixed cost measure captures costs that do not change with the volume of sales; if a firm becomes larger due to random shocks, fixed costs relative to variable costs should decrease.

Finally, we examine the relationship between production concentration and industry output. We use industry real output data from the BEA, available since 1947. In Table 6, we find that industries with higher increases in concentration experience higher growth in real output. Correspondingly, their output shares in the economy increase as well. Figure IA16 visualizes the relationship. Broadly speaking, manufacturing industries witnessed rising concentration and expansion in output in the first half of the 20th century; services and retail/wholesale witnessed rising concentration and expanding output towards the end of the 20th century. One possible concern is that positive idiosyncratic shocks to large firms may lead to both increases in concentration and more industry output in a given period of time. In Table IA4 we also present regressions of changes in concentration over a decade on industry growth in that decade *fitted by industry growth in the past decade*. In this case, we use industry growth in the past decade to capture persistent changes in the industry, and the measurement windows for the left-hand-side and the right-hand-side of the regression are different. The results are similar. Our results also resonate with findings in Ganapati (2021), who analyzes industry-level data from the census in

<sup>21</sup>We do not have direct counterparts of COGS and SG&A in the SOI data. For the cost items in SOI data, it is not easy to assess the extent to which they are variable or fixed.

Table 6 – Rising Concentration and Industry Growth

|                                       | $\Delta_{20}$ Top 1% Asset Share |                     |                     |                     |                     |                     |
|---------------------------------------|----------------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                                       | (1)                              | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
| $\Delta_{20}$ Log Real Gross Output   | 0.068***<br>(0.013)              | 0.051***<br>(0.017) |                     | 0.069***<br>(0.012) | 0.051***<br>(0.016) |                     |
| $\Delta_{20}$ Real Gross Output Share |                                  |                     | 2.297***<br>(0.505) |                     |                     | 2.335***<br>(0.450) |
| Industries                            |                                  | All                 |                     |                     | Nonfinancial        |                     |
| Year Fixed Effect                     | No                               | Yes                 | No                  | No                  | Yes                 | No                  |
| Obs                                   | 1,406                            | 1,406               | 1,406               | 1,304               | 1,304               | 1,304               |
| R <sup>2</sup>                        | 0.06                             | 0.03                | 0.08                | 0.06                | 0.03                | 0.08                |

*Notes:* This table shows industry-level regressions of changes in the top 1% asset share over twenty years on industry growth over twenty years. In columns (1), (2), (4) and (5), industry growth is measured as log changes in real gross output. In columns (3) and (6), industry growth is measured as change in the industry's share in total real gross output of private industries. Columns (1) to (3) show results for all industries. Columns (4) to (6) show results for nonfinancial industries. Year fixed effects are included in the even columns. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in the national accounts follows Tables IA9 and IA10. Standard errors are Driscoll and Kraay (1998) with twenty lags.  $R^2$  does not include fixed effects.

recent decades and documents a positive correlation between changes in CR4 and real output growth.<sup>22</sup>

To formalize how production processes with economies of scale can lead to the empirical facts we observe, we present a simple model in Internet Appendix IA3. In the model, firms with different levels of idiosyncratic productivity can choose between two production technologies: 1) a traditional technology with lower upfront spending but higher marginal costs, and 2) a scalable technology with higher upfront spending but lower marginal costs. On the demand side, we use a standard nested CES structure. We assume an exogenous markup  $\mu$ , which illustrates that the development of the scalable technology does not necessarily increase profitability. The model generates the following results.

1. Firms with very low productivity exit. For the remaining firms, the most productive firms adopt the scalable technology, and the rest adopt the traditional technology.
2. If the development of the scalable technology raises upfront costs and reduces marginal costs, production concentration (e.g., sales shares of the top 1%) would increase. The share of upfront costs in total costs would increase among firms that adopt the scalable technology. The industry's output share in the economy would also increase.
3. Per-period profits/sales remains exogenous; it only depends on the exogenous markup  $\mu$ .

The model illustrates that the development of the scalable technology sets firms further apart. Firms that use the new scalable technology become larger, but not all firms will find it optimal to use this technology given the high upfront spending, so some will stay with the traditional technology and

<sup>22</sup>It is difficult to consistently measure productivity at the industry level over our entire sample period. Productivity measurement may also face other complications (e.g., Solow's paradox, adjustment for product quality, difference between physical productivity and revenue productivity, dependence on assumptions about the production function).

remain small.<sup>23</sup> In this case, the impact of the scalable technology is not necessarily to increase the average firm size, but to “polarize” the size of businesses, as we observe in Figure 4. Meanwhile, the size of total profits will increase for firms that adopt the scalable technology, whereas their profitability (profits/sales) is given by the markup and can be shaped by other forces that affect the markup. In particular, firms that use the scalable technology and incur higher upfront spending are ultimately compensated through greater sales and total profits, not necessarily through a higher per-unit markup. Besides economies of scale, economies of scope can play a role in the formation of large firms as well (Chandler, 1994; Hoberg and Phillips, 2022). To the extent that higher technological intensity also facilitates economies of scope (and fixed costs can be split among both a larger volume and a larger variety of products), our results above are compatible with the economies of scope interpretation.

## 4.2 Trade and Expansion of Markets

If the U.S. consisted of isolated villages, then the size of businesses would face natural limits. Many observers postulate that the initial formation of large U.S. companies in the 19th century was influenced by the integration of domestic markets thanks to railroads, steamships, and telegraphs (Chandler, 1994). It is also natural to ask whether the further expansion of markets in the 20th century, such as globalization, drives the concentration trends in our data. To understand the role of trade and market expansion, we outline a simple theoretical framework in the spirit of Melitz (2003) in Internet Appendix IA3.3. The model makes the following predictions.<sup>24</sup> First, in the absence of heterogeneous technologies (i.e., traditional versus scalable technology), reductions in trade barriers will increase production concentration in the U.S. by allowing top firms to expand by exporting more; however, the concentration of domestic sales (i.e., sales excluding exports) will not change. Second, with heterogeneous technologies, reductions in trade barriers encourage the adoption of the scalable technology, and the concentration of both total sales and domestic sales will increase.

We then discuss the empirical evidence on the role of trade. For international trade, we examine the concentration of domestic sales (“concentration excluding exports”) to understand whether changes in trade barriers alone are sufficient for explaining rising concentration. To be conservative, we subtract all exports from the receipts of the top 1% businesses, and divide this value by total receipts minus exports. Figure IA17 shows that “concentration excluding exports” has increased substantially as well, which suggests that changes in trade barriers may not be the only force affecting rising concentration. Moreover, historical data indicate that barriers to international trade in the US did not decrease much before the 1970s: exports and imports relative to GDP did not expand substantially in the first half of the 20th century, and then they began to increase considerably around 1970 (Wen and Reinbold, 2020). Rising concentration in our data started much earlier. In particular, international trade is especially relevant for manufacturing, but rising concentration in manufacturing was most substantial before the

---

<sup>23</sup>In the simple model, whether a firm finds it appealing to adopt the new versus the old technology is driven by a firm’s productivity. In practice, whether a firm chooses a scalable business model may also depend on the owner’s preferences, managerial capacity, or financing availability.

<sup>24</sup>If the focus is instead product market concentration (rather than production concentration in U.S.), then the relevant metric needs to use sales of U.S. producers minus exports in the numerator and sales of U.S. producers plus imports minus exports as the denominator. See Amiti and Heise (2021) for an analysis of manufacturing products between 1992 and 2012.

1970s. For domestic trade, we cannot directly separate “exports” in the data (i.e., sales to other regions). Nonetheless, prior studies examine the evolution of domestic trade for goods, and find that domestic markets became well-integrated by the late 19th century (Kim, 1995).<sup>25</sup>

In summary, rising concentration does not appear to be driven by changes in trade barriers alone, but having access to large markets could in principle strengthen the influence of economies of scale. In the U.S., broad markets were already available by the early 20th century. Globalization since the 1970s has limitations for explaining the long-run evidence: for manufacturing-related industries, rising concentration largely took place before this era; for services-related industries, rising concentration is stronger in recent decades but the volume of international trade is smaller in services. Some research also suggests that globalization (e.g., Chinese imports) presented a negative shock to large U.S. manufacturers but less so to small ones who produce niche or boutique products (Holmes and Stevens, 2014; Ding et al., 2022), in which case globalization may even reduce production concentration.

### 4.3 Regulations

**Antitrust** Topics related to concentration are frequently mentioned in discussions about antitrust. Syverson (2019) highlights that concentration does not have a clear relationship with competitiveness (higher concentration can be associated with less competition or with more depending on the setting), and De Loecker, Eeckhout, and Unger (2020) emphasize that market power should be measured through markups. Therefore, we do not aim to use our evidence to speak to the strength or weakness of market power, and the success or failure of antitrust policies. However, we can analyze the following question: if we are interested in the role of large firms in overall economic activities, do antitrust policies have a major impact?

Over the past century, antitrust shifted through different regimes; enforcement is generally thought to be tougher before the 1980s and more relaxed afterwards (Peltzman, 2014; Stucke and Ezrachi, 2017; Phillips Sawyer, 2019). Meanwhile, rising concentration occurred throughout these eras. We also analyze the standard measures of antitrust enforcement, such as the annual number of antitrust cases brought by the DOJ (Posner, 1970; Gallo et al., 2000) and the budget of the DOJ antitrust division.<sup>26</sup> Figure IA18 displays these two series, which are about 0.6 correlated. The patterns shown by these two series align with changes in political and judicial philosophies (e.g., an active DOJ after Roosevelt appointed Thurman Arnold to the antitrust division in 1938, and a quieter DOJ with diminished resources in the Reagan era). In Panel A of Table IA5, we examine the average annual DOJ cases and antitrust division budget through presidential cycles. We see that they decrease somewhat when Republicans have presidential or congressional control, and decrease significantly when Republicans control the presidency as well as both chambers of congress. In Panel B, we turn to changes in concentration through these presidential

---

<sup>25</sup>For instance, prices of similar goods converged across regions. Moreover, regional specialization in manufacturing increased in between 1860 and early 1900s but did not increase afterwards. High regional specialization in production suggests that goods were made in centralized locations and shipped nationally.

<sup>26</sup>As Gallo et al. (2000) write, “although DOJ prosecutions provide only a partial picture of all antitrust enforcement effort, omitting FTC, state, and private enforcement efforts, DOJ enforcement efforts constitute an important, if not the dominant, component of American antitrust enforcement.”



cycles (as DOJ resources and activities are affected by the political environment). Overall, we do not observe that changes in corporate concentration display significant relationships with DOJ activities or political environments. One possible concern is that stronger antitrust enforcement or Democratic control can occur in response to more corporate consolidation, which could induce a positive bias in the relationship between DOJ activities and concentration. At a minimum, the long-run data suggest that top business shares have increased through different antitrust regimes. Rising concentration also occurred at different points in time in different industries, and we are not aware that antitrust policies have systematically targeted different industries over time.

Taken together, the data do not show evidence that antitrust is the main determinant of the economy-wide business size distribution throughout the past 100 years. Ultimately, the mandate of antitrust is not necessarily to regulate the economy-wide business size distribution; nonetheless, it may have affected market shares in narrowly defined markets for particular products and locations (Affeldt et al., 2021), which are closer to the realm of antitrust mandates.

**Other regulations** Several other types of regulations may also affect the size of businesses. First, some industry-specific regulations limited the size of business operations during certain time periods, such as restrictions on interstate banking before the 1980s (Savage, 1993) and special taxes on chain stores in the late 1920s and the 1930s in around half of the states (Ross, 1986).<sup>27</sup> Second, a number of policies subsidize small businesses though not explicitly punishing or restricting larger businesses (Hurst and Pugsley, 2011). Third, Akcigit and Ates (2022) suggest that policies and regulations in recent decades limit the diffusion of knowledge, leading to rising concentration and higher markups. Overall, government regulations certainly can affect the size of businesses. To explain our findings, such policies need to have shifted in favor of large firms over the past 100 years. Furthermore, they need to have been particularly important for the growth of large firms in manufacturing rather than in services in the early 20th century, and then switched focus in the later decades. At the moment, we are not aware of such a pattern in regulatory policies.

## 4.4 Other Mechanisms

**Search frictions** Lower search frictions for buyers or higher elasticity of substitution can also raise concentration (Goldmanis et al., 2010; Autor et al., 2020; Albrecht, Menzio, and Vroman, 2022). Measuring changes in search frictions or buyer preferences systematically (across industries and over time) is challenging, and it is less clear why such changes would have affected different industries at different points in time. Future research may find ways to measure and test these mechanisms.

**Demographics** Several papers postulate that the decline of population growth in recent decades has reduced business dynamism and increased the dominance of larger firms (Hopenhayn, Neira, and

---

<sup>27</sup>For the subsector “Finance: Banking”, we do observe that top business shares increased from the 1980s until around the early 2000s (although it can be difficult to establish causation with only one sector-level time series). The movement against retail chains largely ended by the 1940s, but substantial increases in the concentration among retailers only picked up after the 1960s and 1970s. The decline of anti-chain regulations removed obstacles for the development of national retail chains, though it may not be a sufficient condition for the rise of chains.



[Singhania, 2022](#); [Peters and Walsh, 2022](#)). However, the secular decline in population growth only began around the 1960s. In addition, declining population growth alone may not be sufficient to explain the timing of rising concentration across different sectors shown in Section 3. Nonetheless, it is possible that declining population growth raises the general level of concentration in recent decades, even if it is not the full story over the long-run.

**Financing** One might also wonder whether financing availability contributes to the long-run changes in concentration. First, changes in financial frictions may affect the business size distribution. It is possible that financial development over time improves funding availability to smaller businesses, which could decrease concentration. On the other hand, financial development may enhance the adoption of scalable technologies, or help outperforming firms raise funding to stand out from the crowd, which could increase concentration. Empirically, it is well recognized that the degree of financial frictions is difficult to measure. One feature we observe in Figure IA19 is that book equity over assets has declined over the past century (debt financing has increased) among nonfinancial industries (leverage in the financial sector is more affected by regulation), consistent with the secular increase in leverage [Graham, Leary, and Roberts \(2015\)](#) document among listed firms. This feature is present among both the top 1% corporations and the rest, and its timing among different industries does not show the same pattern as the rise in concentration. The data suggest that the availability of external financing could have increased over time for both larger and smaller companies, but it remains difficult to directly test how financial frictions have affected rising concentration. Nonetheless, financial frictions alone do not seem to explain the differences in the timing of rising concentration across different industries (e.g., manufacturing versus services and retail); they probably need to be combined with another mechanism to account for results across different industries. Second, some papers postulate that low interest rates can favor large firms or industry leaders, resulting in higher concentration ([Kroen et al., 2022](#); [Zhao, 2022](#)). This mechanism may play a role in recent decades, but it is not enough for the long-run evidence.

Taken together, increasingly stronger economies of scale appear to align with the long-run evolution of production concentration. However, even if rising production concentration follows from stronger economies of scale, the welfare implications can be nuanced. For example, in the model of [Aghion et al. \(2022\)](#), concentration increases due to lower costs of spanning multiple products or a rising efficiency advantage of large firms. Yet welfare decreases in their baseline parametrization of the model because a new innovator is more likely to face a high-efficiency firm as its competitor, which can discourage innovation and growth in the longer term. [De Loecker, Eeckhout, and Mongey \(2022\)](#) point out that welfare implications are ambiguous if fixed costs of production rise; in this case, the more efficient producers become more dominant, but more resources are tied up in overhead. Finally, [Eeckhout and Veldkamp \(2022\)](#) postulate that certain forces for economies of scale (e.g., data) might increase market power. Accordingly, our evidence sheds light on the long-run changes in the organization of production, but we do not take a stance on whether such developments are necessarily “good” or “socially efficient.”<sup>28</sup>

---

<sup>28</sup>As an extreme example, Marx and Lenin believe that production concentration comes from competition and technology-driven economies of scale, but the dominance of large private businesses is undesirable (and should be replaced by central planning to take advantage of economies of scale while avoiding abuse by capitalists). [Marx \(1867\)](#) wrote: “The battle of competition is fought by cheapening of commodities. The cheapness of commodities demands, *caeteris paribus*, on the

## 5 Conclusion

We collect long-run data on the size distribution of U.S. corporations, and document that corporate concentration has been rising over the past century. The rise was stronger in manufacturing and mining in earlier decades, and stronger in services, retail, and wholesale in later decades. We find that the timing and the degree of rising concentration in an industry align closely with the intensity of IT and R&D and the prevalence of fixed costs. Industries with higher increases in concentration also exhibit higher output growth. Among the leading hypotheses for rising concentration, stronger economies of scale appear most consistent with the long-run trends. Our findings on the evolution of production concentration also have implications for the determinants of macroeconomic outcomes, such as the effects of shocks to larger versus smaller firms (Gabaix, 2011) and the effects of financial frictions across the firm size distribution (Crouzet and Mehrotra, 2020).

Although the long-run trends align with stronger economies of scale, our results do not rule out that some large firms may have expanded their territories by unduly exerting power and influence (Cunningham, Ederer, and Ma, 2021; Kamepalli, Rajan, and Zingales, 2022). It is also possible that regulatory policies or business environments have been more favorable to larger companies in recent years (Philippon, 2019; Kroen et al., 2022); special features of a given era can coexist with long-run forces. Finally, the welfare consequences of rising concentration are not obvious even if this development comes from economies of scale. As Stigler (1958) writes in his article on “The Economies of Scale,” “the socially optimum firm is fundamentally an ethical concept, and we question neither its importance nor its elusiveness.”

An intriguing question is whether rising concentration will be an enduring trend in the future. One might ask why technological forces in the past century have shifted toward stronger economies of scale. A possibility is that businesses always seek to expand, and technology has been enhancing the replication of production processes as Brynjolfsson et al. (2008) highlight; moreover, such replication is easier to implement within a firm. Will such forces dominate perpetually? The answer is not obvious. At the turn of the 20th century, the *inevitability* of technological changes leading to increasingly larger enterprises and higher production concentration was a central doctrine of the communists. Lenin, for example, believed economies of scale due to “modern technology” would be so strong that the Soviet Union could be run by one giant firm to enhance efficiency. Such an extreme view perhaps remains unrealistic, but this exact proposition inspired Coase (1937)’s inquiry about the boundaries of the firm, as explained in Coase (1988). Indeed, discussions about production concentration during that era shaped a number of prominent intellectual traditions: some maintain that large enterprises would become all powerful and change the way the society should be organized, whereas others caution that large organizations face certain limitations (Berle and Means, 1932; Lange, 1937; Schumpeter, 1942; Hayek, 1945). More analyses about the nature of the firm and the foundations for the organization of production may provide knowledge that can guide our outlook for the future.

---

productiveness of labour, and this again on the scale of production. Therefore, the larger capitals beat the smaller.” In Lenin’s summary: “the victory of large-scale production is immediately apparent...and the peasant economy...declines and falls into ruin under the burden of its backward technique...By destroying small-scale production, capital leads to an increase in productivity of labour...but the product of this collective labour is appropriated by a handful of capitalists.”

## References

- Adelman, Morris Albert.** The measurement of industrial concentration. *Review of Economics and Statistics*, 1951. 33(4):269–296.
- Affeldt, Pauline, Tomaso Duso, Klaus Peter Gugler, and Joanna Piechucka.** Market concentration in Europe: Evidence from antitrust markets. Working paper, 2021.
- Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li.** A theory of falling growth and rising rents. *Review of Economic Studies*, 2022. Forthcoming.
- Akcigit, Ufuk and Sina T. Ates.** What happened to US business dynamism? *Journal of Political Economy*, 2022. Forthcoming.
- Albrecht, James, Guido Menzio, and Susan Vroman.** Vertical differentiation in frictional product markets. Working paper, 2022.
- Ali, Ashiq, Sandy Klasa, and Eric Yeung.** The limitations of industry concentration measures constructed with Compustat data: Implications for finance research. *Review of Financial Studies*, 2008. 22(10):3839–3871.
- Amiti, Mary and Sebastian Heise.** US market concentration and import competition. Working paper, 2021.
- Anderson, Mark C., Rajiv D. Banker, and Surya N. Janakiraman.** Are selling, general, and administrative costs “sticky”? *Journal of Accounting Research*, 2003. 41(1):47–63.
- Auerbach, Alan J.** The 2021 Martin Feldstein lecture: The taxation of business income in the global economy. [www.nber.org/lecture/2021-martin-feldstein-lecture-taxation-business-income-global-economy](http://www.nber.org/lecture/2021-martin-feldstein-lecture-taxation-business-income-global-economy), 2021. NBER.
- Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen.** The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics*, 2020. 135(2):645–709.
- Backus, Matthew, Christopher Conlon, and Michael Sinkinson.** The common ownership hypothesis: Theory and evidence. Working paper, 2019.
- Barkai, Simcha.** Declining labor and capital shares. *Journal of Finance*, 2020. 75(5):2421–2463.
- Barkai, Simcha and Seth G. Benzell.** 70 years of US corporate profits. Working paper, 2018.
- Benkard, C. Lanier, Ali Yurukoglu, and Anthony Lee Zhang.** Concentration in product markets. Working paper, 2021.
- Berle, Adolf and Gardiner Means.** *The Modern Corporation and Private Property*. Harcourt, Brace & Company, 1932.
- Bessen, James.** Industry concentration and information technology. *Journal of Law and Economics*, 2020. 63(3):531–555.
- Blair, John Malcolm, Harrison F. Houghton, and Matthew Rose.** *Economic Concentration and World War II: Report of the Smaller War Plants Corporation to the Special Committee to Study Problems of American Small Business, United States Senate*. US Government Printing Office, 1946.
- Blanchet, Thomas, Juliette Fournier, and Thomas Piketty.** Generalized Pareto curves: Theory and applications. *Review of Income and Wealth*, 2022. 68(1):263–288.

- Brynjolfsson, Erik, Andrew McAfee, Michael Sorell, and Feng Zhu.** Scale without mass: Business process replication and industry dynamics. Working paper, 2008.
- Chandler, Alfred D.** *Scale and Scope: The Dynamics of Industrial Capitalism*. Harvard University Press, 1994.
- Clarke, Conor and Wojciech Kopczuk.** Business income and business taxation in the United States since the 1950s. *Tax Policy and the Economy*, 2017. 31(1):121–159.
- Coase, Ronald H.** The nature of the firm. *Economica*, 1937. 4(16):386–405.
- Coase, Ronald H.** The nature of the firm: Origin. *Journal of Law, Economics, & Organization*, 1988. 4(1):3–17.
- Collins, Norman R. and Lee E. Preston.** The size structure of the largest industrial firms, 1909–1958. *American Economic Review*, 1961. 51(5):986–1011.
- Committee on Recent Economic Changes.** *Recent Economic Changes in the United States*. NBER, 1929.
- Covarrubias, Matias, Germán Gutiérrez, and Thomas Philippon.** From good to bad concentration? US industries over the past 30 years. *NBER Macroeconomics Annual*, 2020. 34(1):1–46.
- Crouzet, Nicolas and Janice Eberly.** Understanding weak capital investment: The role of market concentration and intangibles. Jackson Hole Symposium, 2019.
- Crouzet, Nicolas and Janice Eberly.** Rents and intangible capital: A q+ framework. *Journal of Finance*, 2021. Forthcoming.
- Crouzet, Nicolas and Neil R. Mehrotra.** Small and large firms over the business cycle. *American Economic Review*, 2020. 110(11):3549–3601.
- Cunningham, Colleen, Florian Ederer, and Song Ma.** Killer acquisitions. *Journal of Political Economy*, 2021. 129(3):649–702.
- De Loecker, Jan, Jan Eeckhout, and Simon Mongey.** Quantifying market power and business dynamism. Working paper, 2022.
- De Loecker, Jan, Jan Eeckhout, and Gabriel Unger.** The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, 2020. 135(2):561–644.
- DeAngelo, Harry, Linda DeAngelo, and Douglas J. Skinner.** Are dividends disappearing? Dividend concentration and the consolidation of earnings. *Journal of Financial Economics*, 2004. 72(3):425–456.
- Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda.** The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives*, 2014a. 28(3):3–24.
- Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda.** The secular decline in business dynamism in the US. Working paper, 2014b.
- Demsetz, Harold.** Industry structure, market rivalry, and public policy. *Journal of Law and Economics*, 1973. 16(1):1–9.
- Ding, Xiang, Teresa C. Fort, Stephen J. Redding, and Peter K. Schott.** Structural change within versus across firms: Evidence from the United States. Working paper, 2022.
- Driscoll, John C. and Aart C. Kraay.** Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 1998. 80(4):549–560.

- Eckert, Fabian, Teresa C. Fort, Peter K. Schott, and Natalie J. Yang.** Imputing missing values in the US Census Bureau's County Business Patterns. Working paper, 2020.
- Eeckhout, Jan.** *The Profit Paradox: How Thriving Firms Threaten the Future of Work*. Princeton University Press, 2021.
- Eeckhout, Jan and Laura Veldkamp.** Data and market power. Working paper, 2022.
- Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin.** The decline of the US labor share. *Brookings Papers on Economic Activity*, 2013. 2013(2):1–63.
- Firooz, Hamid, Zheng Liu, and Yajie Wang.** Automation, market concentration, and the labor share. Working paper, 2022.
- Foster, Lucia, John Haltiwanger, and Cody Tuttle.** Rising markups or changing technology. Working paper, 2022.
- Frank, Robert H. and Philip J. Cook.** *The Winner-Take-All Society: Why the Few at the Top Get So Much More Than the Rest of Us*. Penguin Books, 1996.
- Frydman, Carola and Raven E. Saks.** Executive compensation: A new view from a long-term perspective, 1936–2005. *Review of Financial Studies*, 2010. 23(5):2099–2138.
- Gabaix, Xavier.** The granular origins of aggregate fluctuations. *Econometrica*, 2011. 79(3):733–772.
- Gallo, Joseph C., Kenneth Dau-Schmidt, Joseph L. Craycraft, and Charles J. Parker.** Department of Justice antitrust enforcement, 1955–1997: An empirical study. *Review of Industrial Organization*, 2000. 17(1):75–133.
- Ganapati, Sharat.** Growing oligopolies, prices, output, and productivity. *American Economic Journal: Microeconomics*, 2021. 13(3):309–27.
- Goldmanis, Maris, Ali Hortaçsu, Chad Syverson, and Önsel Emre.** E-commerce and the market structure of retail industries. *Economic Journal*, 2010. 120(545):651–682.
- Graham, John R., Mark T. Leary, and Michael R. Roberts.** A century of capital structure: The leveraging of corporate America. *Journal of Financial Economics*, 2015. 118(3):658–683.
- Grullon, Gustavo, Yelena Larkin, and Roni Michaely.** Are US industries becoming more concentrated? *Review of Finance*, 2019. 23(4):697–743.
- Gutiérrez, Germán and Thomas Philippon.** Investmentless growth: An empirical investigation. *Brookings Papers on Economic Activity*, 2017. 2017(2):89–190.
- Gutiérrez, Germán and Thomas Philippon.** Some facts about dominant firms. Working paper, 2020.
- Gutiérrez, Germán and Thomas Philippon.** How European markets became free: A study of institutional drift. *Journal of the European Economic Association*, 2022. page jvac071.
- Hall, Robert E.** New evidence on the markup of prices over marginal costs and the role of mega-firms in the US economy. Working paper, 2018.
- Harris, Marty and Ken Szeftlinski.** Celebrating ninety years of SOI: Selected corporate data, 1916–2004. *SOI Bulletin*, 2007. 27(2):279–291.
- Haskel, Jonathan and Stian Westlake.** *Capitalism without Capital*. Princeton University Press, 2017.
- Hayek, Frederick August von.** The use of knowledge in society. *American Economic Review*, 1945. 35(4):519–530.

- Hoberg, Gerard and Gordon M. Phillips.** Scope, scale and competition: The 21st century firm. Working paper, 2022.
- Holmes, Thomas J. and John J. Stevens.** An alternative theory of the plant size distribution, with geography and intra-and international trade. *Journal of Political Economy*, 2014. 122(2):369–421.
- Holmstrom, Bengt and Steven N. Kaplan.** Corporate governance and merger activity in the united states: Making sense of the 1980s and 1990s. *Journal of Economic Perspectives*, 2001. 15(2):121–144.
- Hopenhayn, Hugo, Julian Neira, and Rish Singhania.** From population growth to firm demographics: Implications for concentration, entrepreneurship and the labor share. *Econometrica*, 2022. 90(4):1879–1914.
- Hsieh, Chang-Tai and Esteban Rossi-Hansberg.** The industrial revolution in services. *Journal of Political Economy: Macroeconomics*, 2022. Forthcoming.
- Hubmer, Joachim.** The race between preferences and technology. *Econometrica*, 2022. Forthcoming.
- Hubmer, Joachim and Pascual Restrepo.** Not a typical firm: Capital-labor substitution and firms' labor shares. Working paper, 2022.
- Hurst, Erik and Benjamin Wild Pugsley.** What do small businesses do? *Brookings Papers on Economic Activity*, 2011. pages 73–143.
- Kamepalli, Sai Krishna, Raghuram Rajan, and Luigi Zingales.** Kill zone. Working paper, 2022.
- Kaplan, Steven N. and Joshua Rauh.** It's the market: The broad-based rise in the return to top talent. *Journal of Economic Perspectives*, 2013. 27(3):35–56.
- Keil, Jan.** The trouble with approximating industry concentration from compustat. *Journal of Corporate Finance*, 2017. 45:467–479.
- Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy.** Measuring technological innovation over the long run. *American Economic Review: Insights*, 2021. 3(3):303–20.
- Kim, Sukkoo.** Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860–1987. *Quarterly Journal of Economics*, 1995. 110(4):881–908.
- Kondo, Illenin, Logan T. Lewis, and Andrea Stella.** On the US firm and establishment size distributions. Working paper, 2019.
- Kopczuk, Wojciech and Eric Zwick.** Business incomes at the top. *Journal of Economic Perspectives*, 2020. 34(4):27–51.
- Kroen, Thomas, Ernest Liu, Atif R. Mian, and Amir Sufi.** Falling rates and rising superstars. Working paper, 2022.
- Kuhn, Moritz, Moritz Schularick, and Ulrike I. Steins.** Income and wealth inequality in America. *Journal of Political Economy*, 2020. 128(9):3469–3519.
- Lamoreaux, Naomi R.** Active proprietorships, partnerships, and corporations – entities, receipts, and profits. In Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, editors, *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition*. Cambridge University Press, New York, 2006.
- Lamoreaux, Naomi R.** The problem of bigness: From Standard Oil to Google. *Journal of Economic Perspectives*, 2019. 33(3):94–117.



- Lange, Oskar.** On the economic theory of socialism: Part two. *Review of Economic Studies*, 1937. 4(2):123–142.
- Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard.** Information technology and returns to scale. Working paper, 2022.
- Lenin, Vladimir Ilich.** *Imperialism: The Highest Stage of Capitalism*. Life and Knowledge Publishers, 1916.
- Luttmer, Erzo G. J.** Models of growth and firm heterogeneity. *Annual Review of Economics*, 2010. 2(1):547–576.
- Marshall, Alfred.** *Principles of Economics*. Macmillan, 1890.
- Marx, Karl.** *Das Kapital*, volume 1. Otto Meissner, 1867.
- Means, Gardiner C.** The growth in the relative importance of the large corporation in American economic life. *American Economic Review*, 1931. pages 10–42.
- Means, Gardiner Coit, Donald Ewan Montgomery, John Maurice Clark, Alvin Harvey Hansen, and Mordecai Ezekiel.** *The Structure of the American Economy*, volume 1. US Government Printing Office, 1939.
- Melitz, Marc J.** The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 2003. 71(6):1695–1725.
- Mills, Lillian F., Kaye J. Newberry, and William B Trautman.** Trends in book-tax income and balance sheet differences. Working paper, 2002.
- Peltzman, Sam.** Industrial concentration under the rule of reason. *Journal of Law and Economics*, 2014. 57(S3):S101–S120.
- Persky, Joseph.** Retrospectives: Lange and von Mises, large-scale enterprises, and the economic case for socialism. *Journal of Economic Perspectives*, 1991. 5(4):229–236.
- Peters, Michael and Conor Walsh.** Population growth and firm-product dynamics. Working paper, 2022.
- Petska, Thomas B. and Robert A. Wilson.** Trends in business structure and activity, 1980-1990. *SOI Bulletin*, 1994. 13:27–72.
- Philippon, Thomas.** *The Great Reversal: How America Gave Up on Free Markets*. Harvard University Press, 2019.
- Phillips Sawyer, Laura.** US antitrust law and policy in historical perspective. In *Oxford Research Encyclopedia of American History*. 2019.
- Piketty, Thomas and Emmanuel Saez.** Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 2003. 118(1):1–41.
- Posner, Richard A.** A statistical study of antitrust enforcement. *Journal of Law and Economics*, 1970. 13(2):365–419.
- Pryor, Frederic L.** New trends in US industrial concentration. *Review of Industrial Organization*, 2001. 18(3):301–326.
- Rosen, Sherwin.** The economics of superstars. *American Economic Review*, 1981. 71(5):845–858.
- Ross, Thomas W.** Store wars: The chain tax movement. *Journal of Law and Economics*, 1986. 29(1):125–137.

- Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter.** Diverging trends in national and local concentration. *NBER Macroeconomics Annual*, 2021. 35(1):115–150.
- Rossi-Hansberg, Esteban and Mark L. J. Wright.** Establishment size dynamics in the aggregate economy. *American Economic Review*, 2007. 97(5):1639–1666.
- Saez, Emmanuel and Gabriel Zucman.** Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 2016. 131(2):519–578.
- Savage, Donald T.** Interstate banking: A status report. *Federal Reserve Bulletin*, 1993. 79:1075.
- Schmalz, Martin C.** Common-ownership concentration and corporate conduct. *Annual Review of Financial Economics*, 2018. 10.
- Schumpeter, Joseph Alois.** *Capitalism, Socialism and Democracy*. Allen & Unwin London, 1942.
- Singla, Shikhar.** Regulatory costs and market power. Working paper, 2022.
- Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick.** Capitalists in the twenty-first century. *Quarterly Journal of Economics*, 2019. 134(4):1675–1745.
- Stigler, George J.** The economies of scale. *Journal of Law and Economics*, 1958. 1:54–71.
- Streuling, Guenther Fred.** *A Comparison of Income Tax Regulations for Consolidated Returns and Generally Accepted Consolidation Practices for Financial Statements*. 1971.
- Stucke, Maurice E. and Ariel Ezrachi.** The rise, fall, and rebirth of the US antitrust movement. *Harvard Business Review*, 2017. 15.
- Sutton, John.** *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. MIT press, 1991.
- Sutton, John.** Gibrat’s legacy. *Journal of Economic Literature*, 1997. 35(1):40–59.
- Sutton, John.** *Technology and Market Structure: Theory and History*. MIT press, 2001.
- Syverson, Chad.** Macroeconomics and market power: Context, implications, and open questions. *Journal of Economic Perspectives*, 2019. 33(3):23–43.
- Traina, James.** Is aggregate market power increasing? Production trends using financial statements. Working paper, 2018.
- Wen, Yi and Brian Reinbold.** The evolution of total trade in the U.S. <https://www.stlouisfed.org/on-the-economy/2020/march/evolution-total-trade-us>, 2020. Federal Reserve Bank of St. Louis.
- White, Lawrence J.** Trends in aggregate concentration in the United States. *Journal of Economic Perspectives*, 2002. 16(4):137–160.
- Zhao, Dongchen.** Low interest rates, debt heterogeneity, and firm dynamics. Working paper, 2022.

## Internet Appendix: For Online Publication

### IA1 Additional Figures and Tables

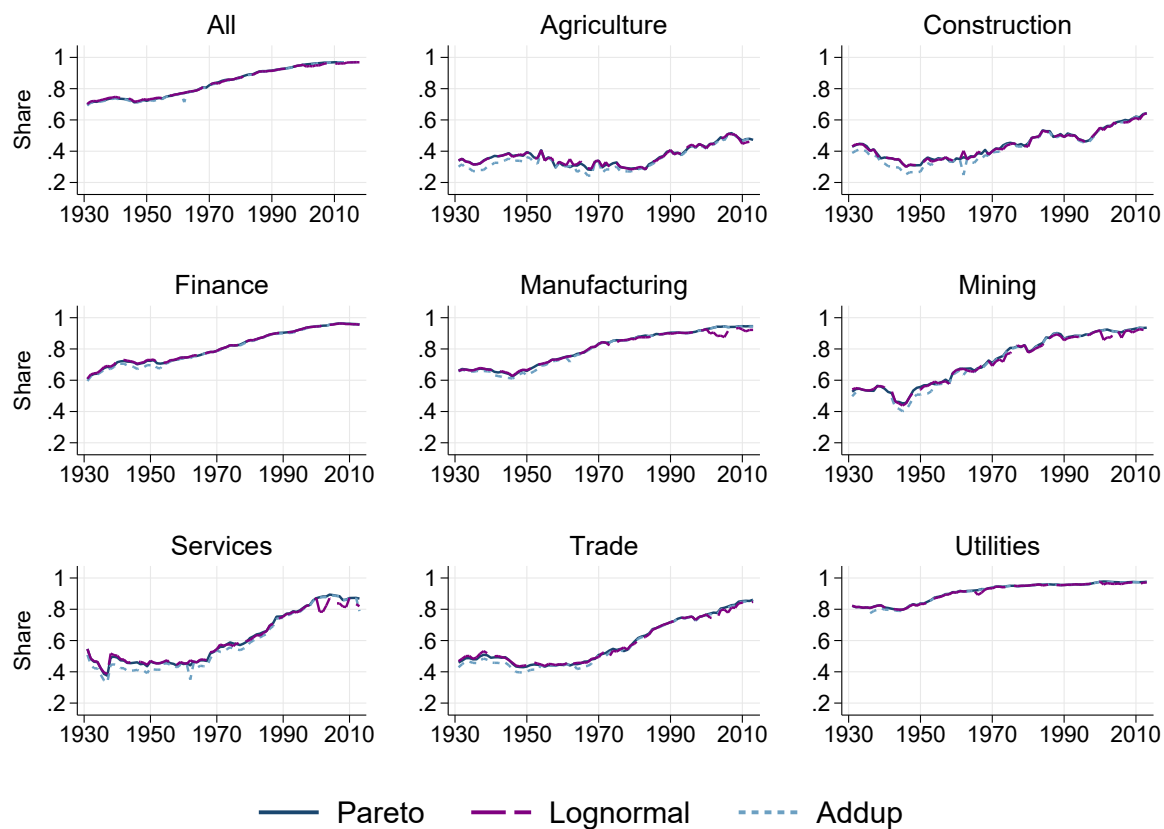


Figure IA1. Comparison of Different Methods for Estimating Top Shares

*Notes:* This figure shows the top 1% asset share using the three methods explained in Section 2. The solid line shows the results of interpolating Pareto distributions. The dashed line shows the results of interpolating lognormal distributions. The dotted line shows the results of adding up top bins.

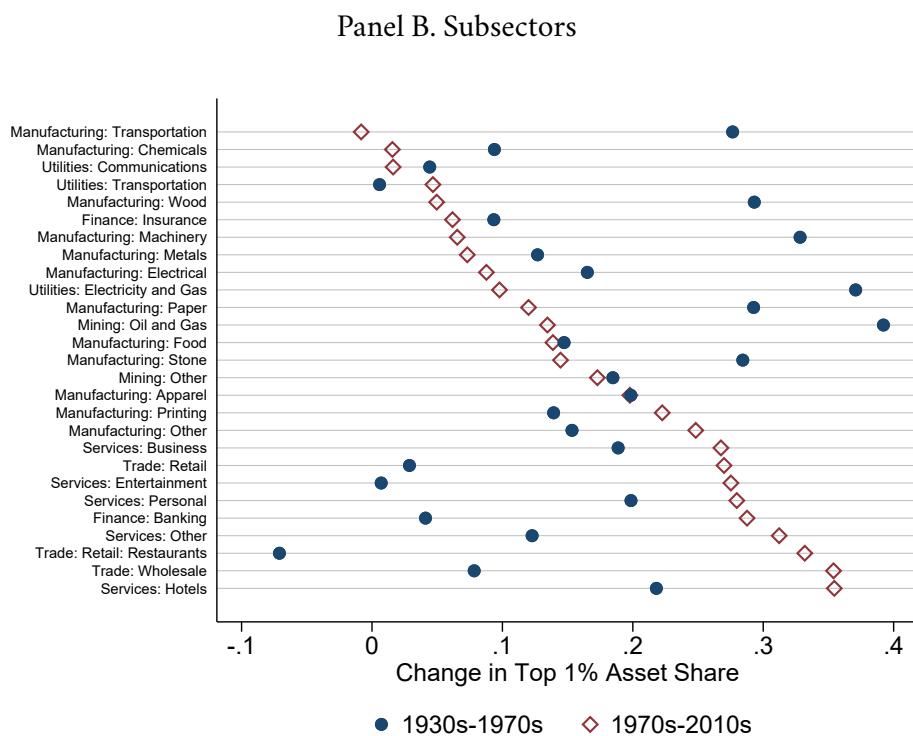
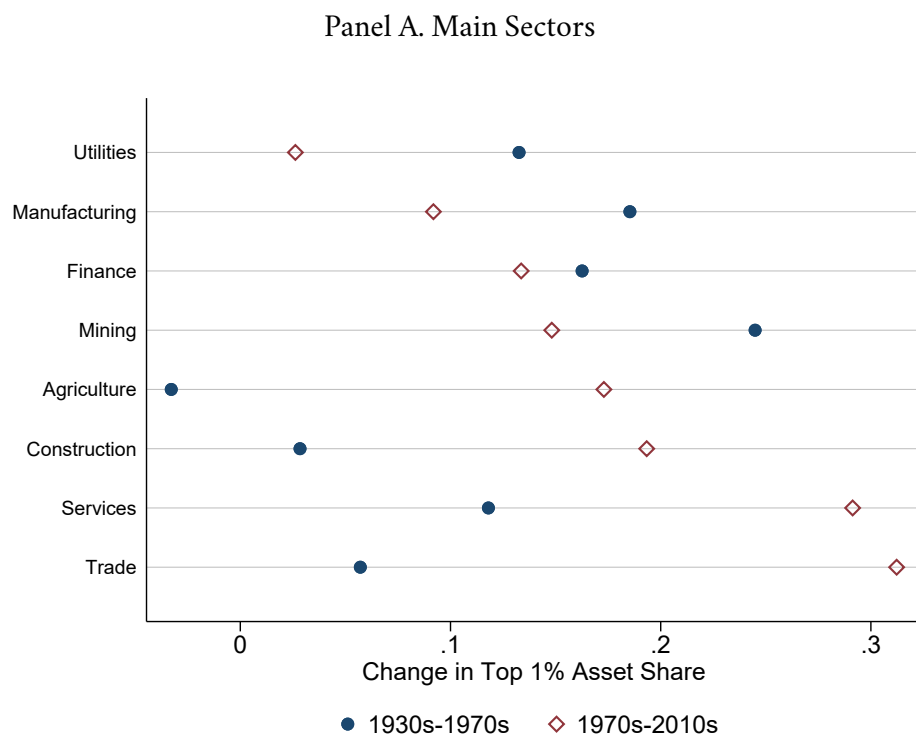
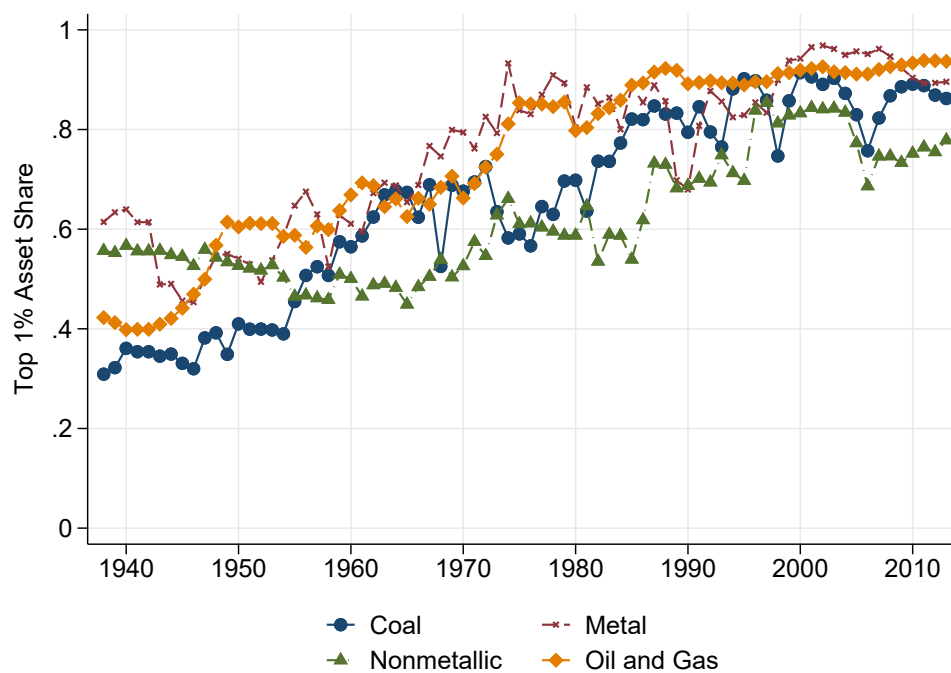


Figure IA2. Rising Concentration in Earlier and Later Decades

Notes: This figure shows the change in the top 1% asset share between the 1930s to the 1970s (solid blue circle) and between the 1970s and the 2010s (hollow red diamond), for main sectors in Panel A and subsectors in Panel B. The industries are sorted by the change in the top 1% asset share between the 1970s and the 2010s.

Panel A. Mining



Panel B. Retail

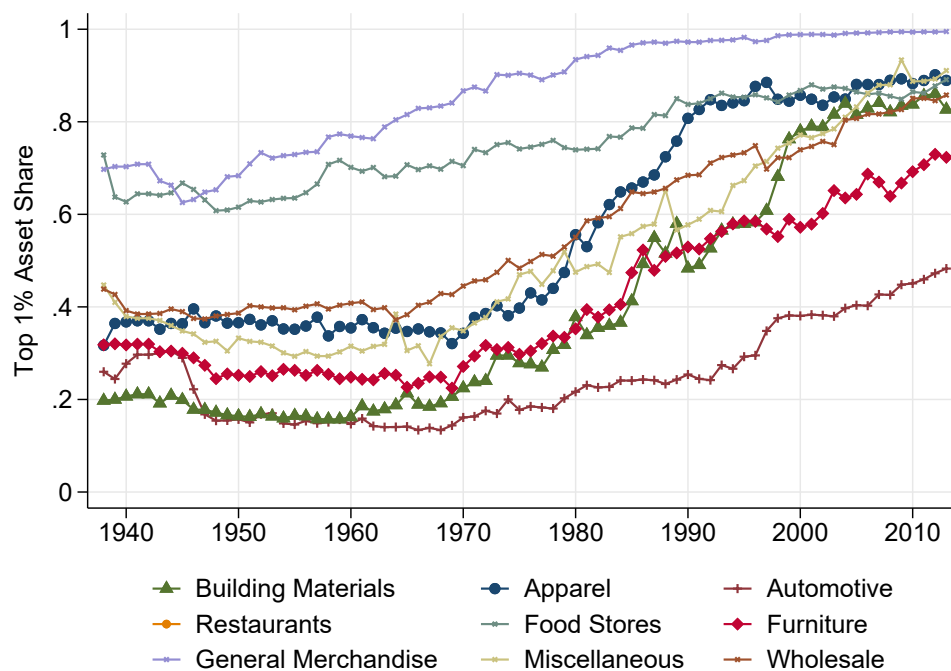


Figure IA3. Concentration within Subsectors

Notes: This figure shows further breakdowns of mining subsectors in Panel A and retail subsectors in Panel B. It shows the top 1% asset share in each segment of mining and retail.

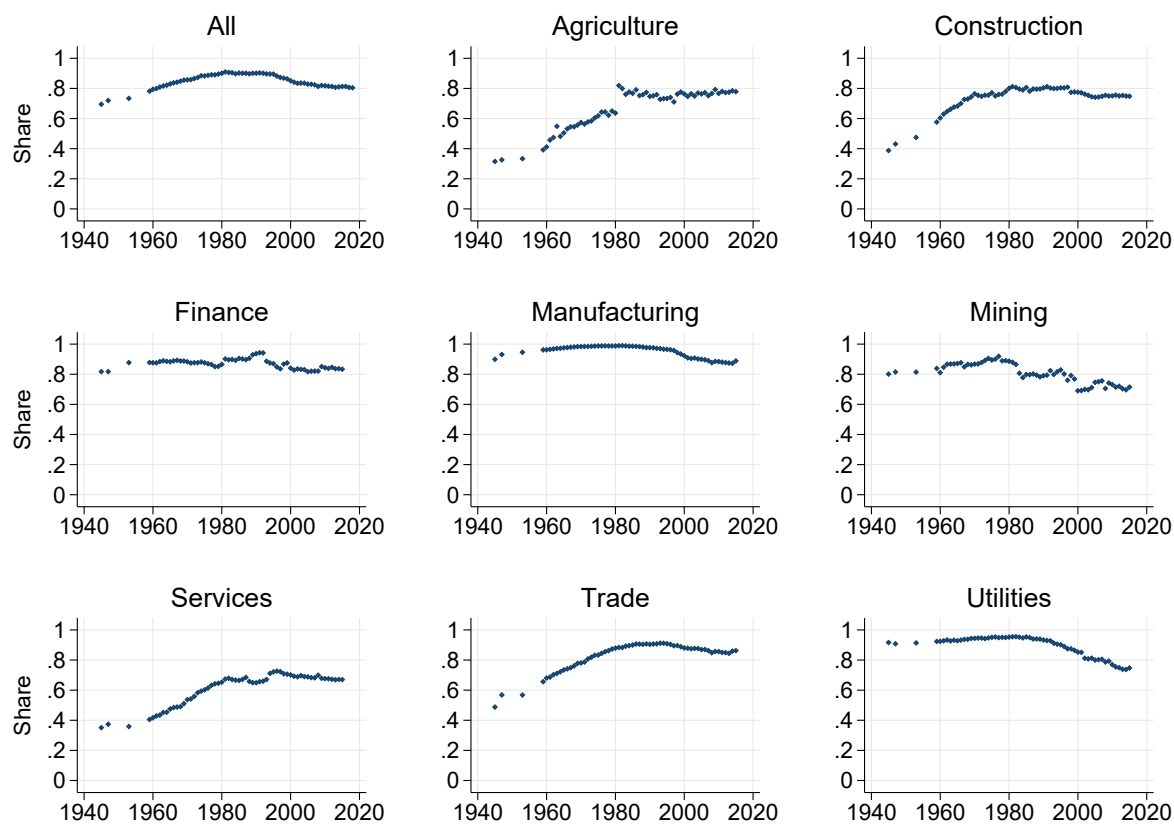


Figure IA4. Corporations in Total Business Receipts

*Notes:* This figure shows the share of corporations in the total value of business receipts (receipts by corporations, partnerships, and nonfarm proprietorships), for the aggregate and the main sectors.



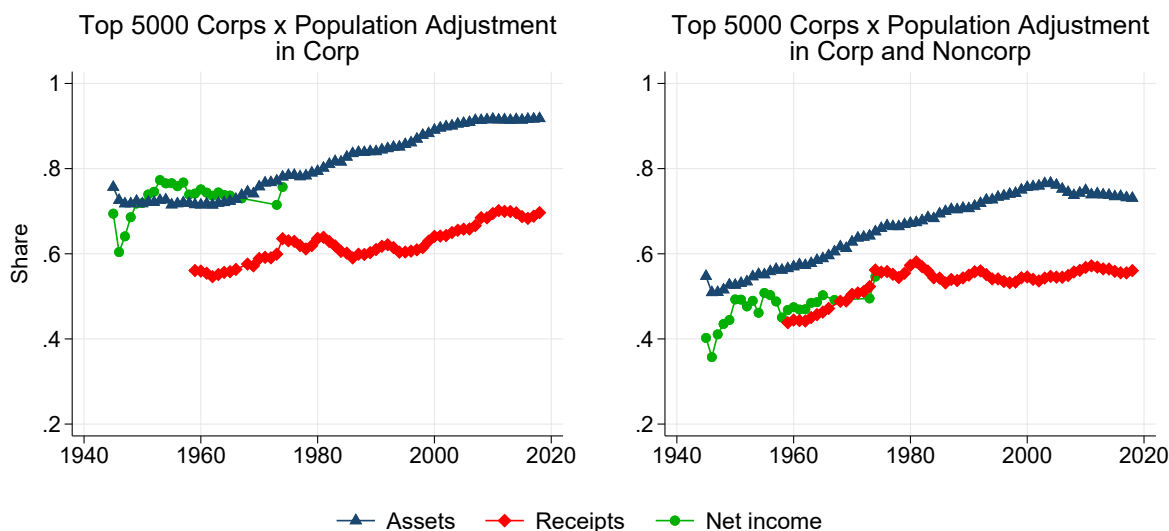
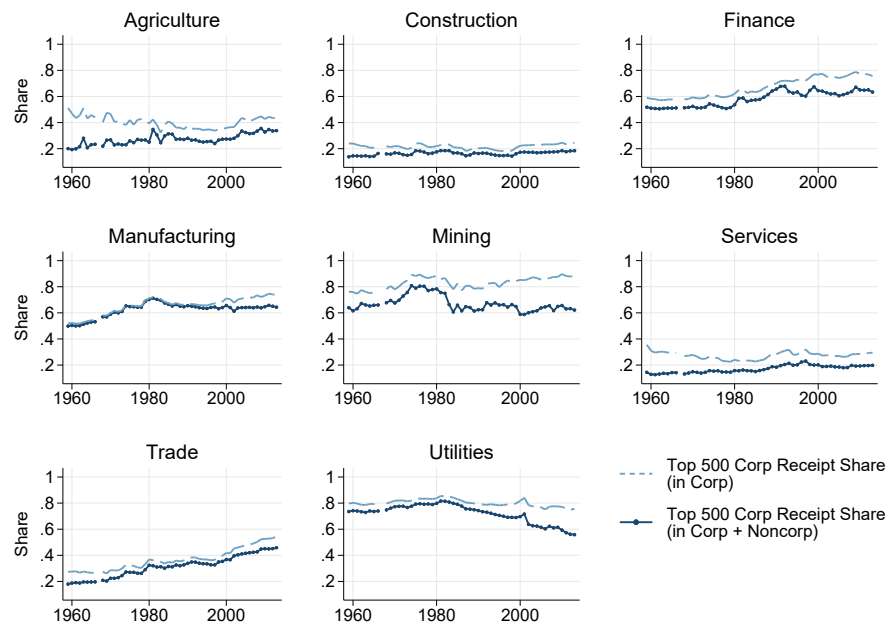


Figure IA5. Shares of Top Corporations: Population Growth Adjustment

*Notes:* This figure shows top shares when the number of top firms increases with population growth. The numerator includes the top 5,000 corporations in 1945, and the number of top corporations increases based on the cumulative population growth afterwards. The denominator is based on all corporations (left panel) and all corporations plus noncorporations (right panel). The blue line with triangles shows the share of assets accounted for by top corporations sorted on assets. The red line with diamonds shows the share of receipts accounted for by top corporations sorted on receipts. The green line with circles shows the share of net income accounted for by top corporations sorted on net income (restricting to those with positive net income). See Internet Appendix [IA2.3](#) for details about variable construction.

### Panel A. Baseline



### Panel B. Population Growth Adjustment

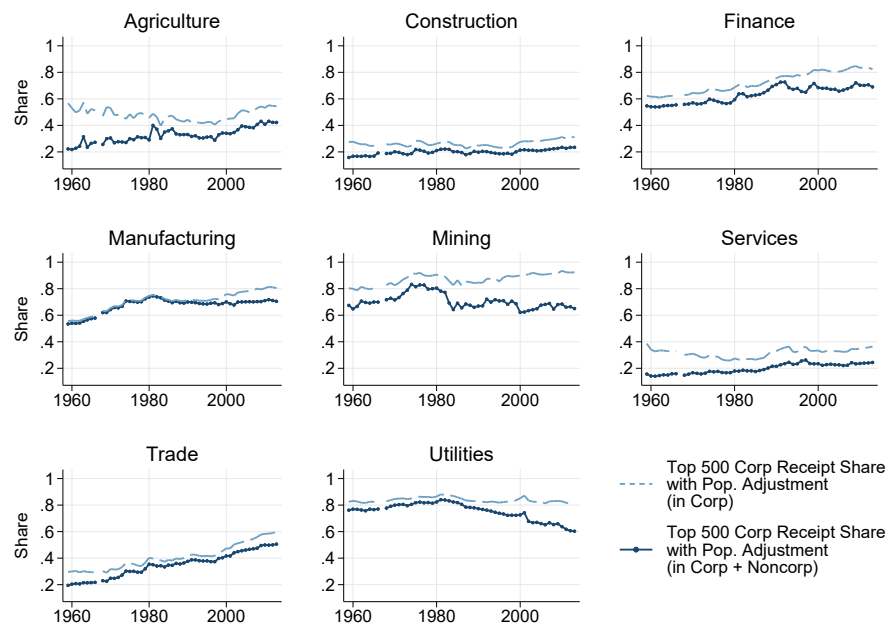
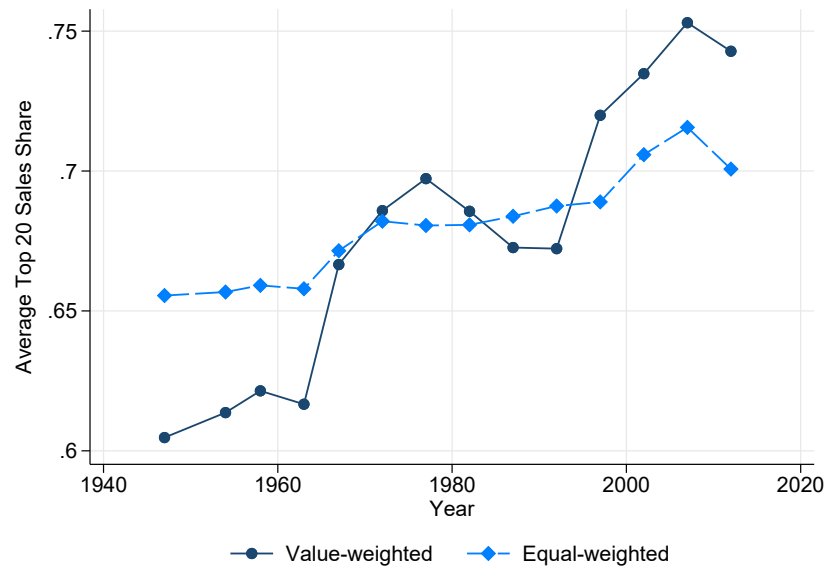


Figure IA6. Receipt Shares of Top 500 Corporations: Main Sectors

Notes: Panel A shows the share of the top 500 corporations by receipts in the total receipts of all corporations (dashed line) and the total receipts of corporations plus noncorporations (solid line). In Panel B we start with the receipt share of top 500 corporations at the beginning of the sample, and increase the number of top corporations in the numerator using the cumulative population growth afterwards. The denominator is based on the total receipts of all corporations (dashed line) and the total receipts of corporations plus noncorporations (solid line). See Internet Appendix [IA2.3](#) for details about variable construction.

Panel A. Time Series of Top 20 Share in Manufacturing Census



Panel B. Cross Section of Top 20 Share in 2012 Census

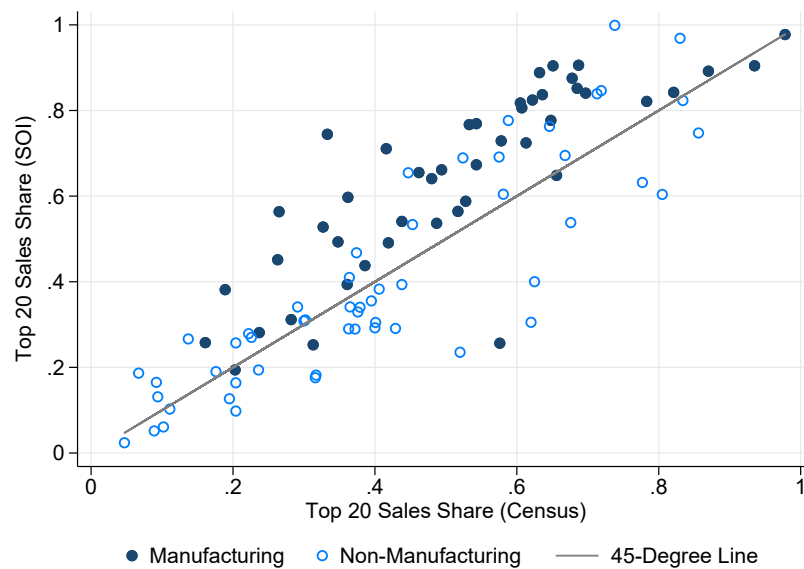
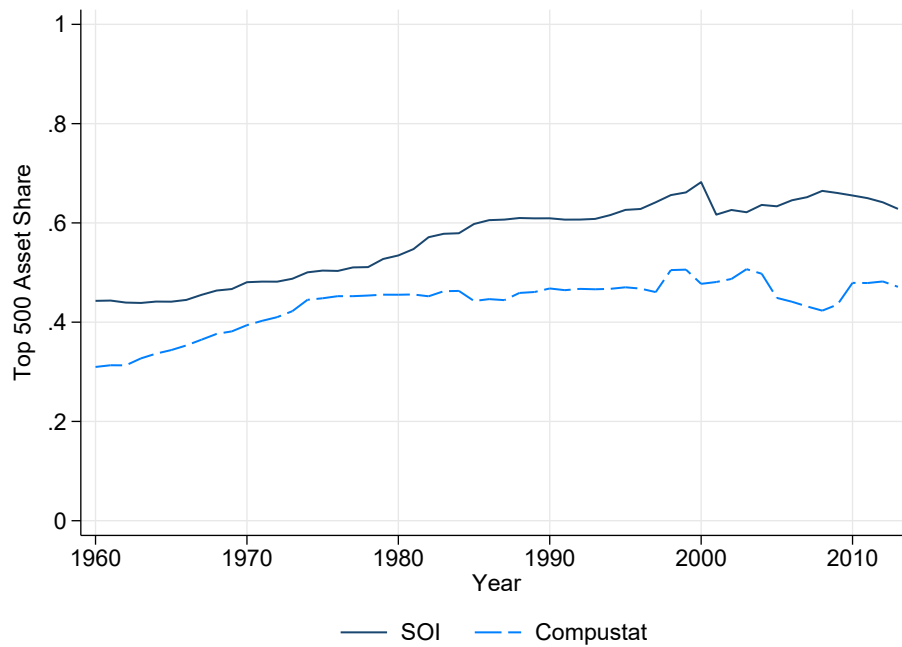


Figure IA7. Concentration Ratios in Census Data

*Notes:* Panel A shows the time series of the value-weighted average (solid line with circles) and equal-weighted average (dashed line with diamonds) of sales concentration of the top 20 firms by sales from the manufacturing census. The data use four-digit SIC industries until 1992 and six-digit NAICS industries after 1997. Panel B shows the 2012 cross section of the sales share of the top 20 firms by sales in census data on the  $x$ -axis, and sales share of the top 20 businesses by assets in SOI data on the  $y$ -axis. Each dot is a six-digit SOI industry, which largely map into four-digit NAICS. Solid dots indicate manufacturing industries and hollow dots indicate non-manufacturing industries. See Internet Appendix [IA2.3](#) for details about variable construction.

Panel A. Top 500 Share by Assets in Total Corporate Assets



Panel B. Top 500 Share by Receipts in Total Corporate Receipts

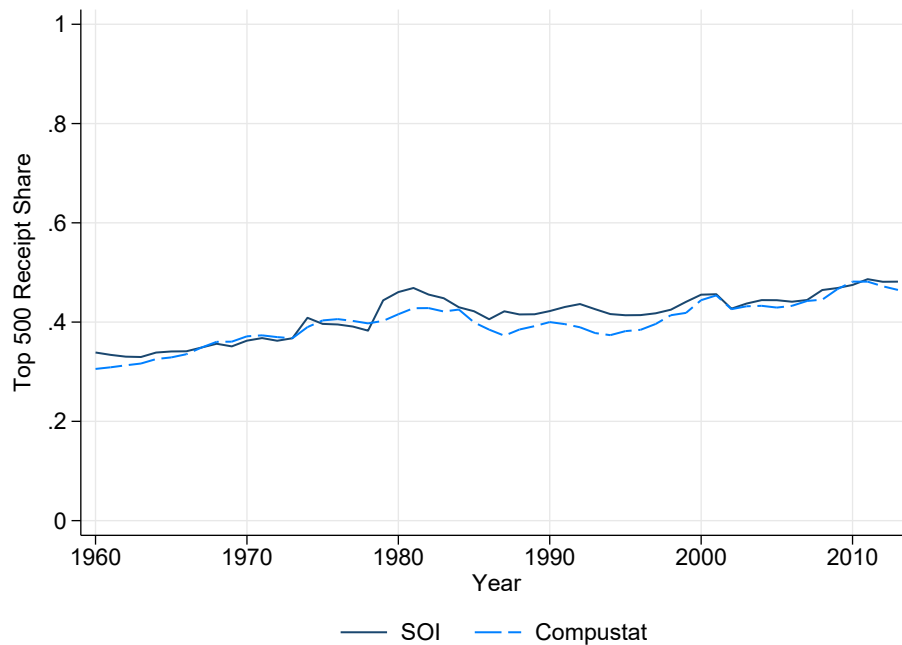


Figure IA8. Comparison with Compustat

*Notes:* This figure compares the share of the top 500 businesses estimated from SOI and Compustat data. Panel A shows the imputed share of the top 500 corporations by assets in total corporate assets using SOI data (solid line), and the share of the top 500 by assets in Compustat in total corporate assets (dashed line). Panel B shows the imputed share of the top 500 corporations by receipts in total corporate receipts using SOI data (solid line), and the share of the top 500 by sales in Compustat in total corporate receipts (dashed line). See Internet Appendix [IA2.3](#) for details about variable construction.

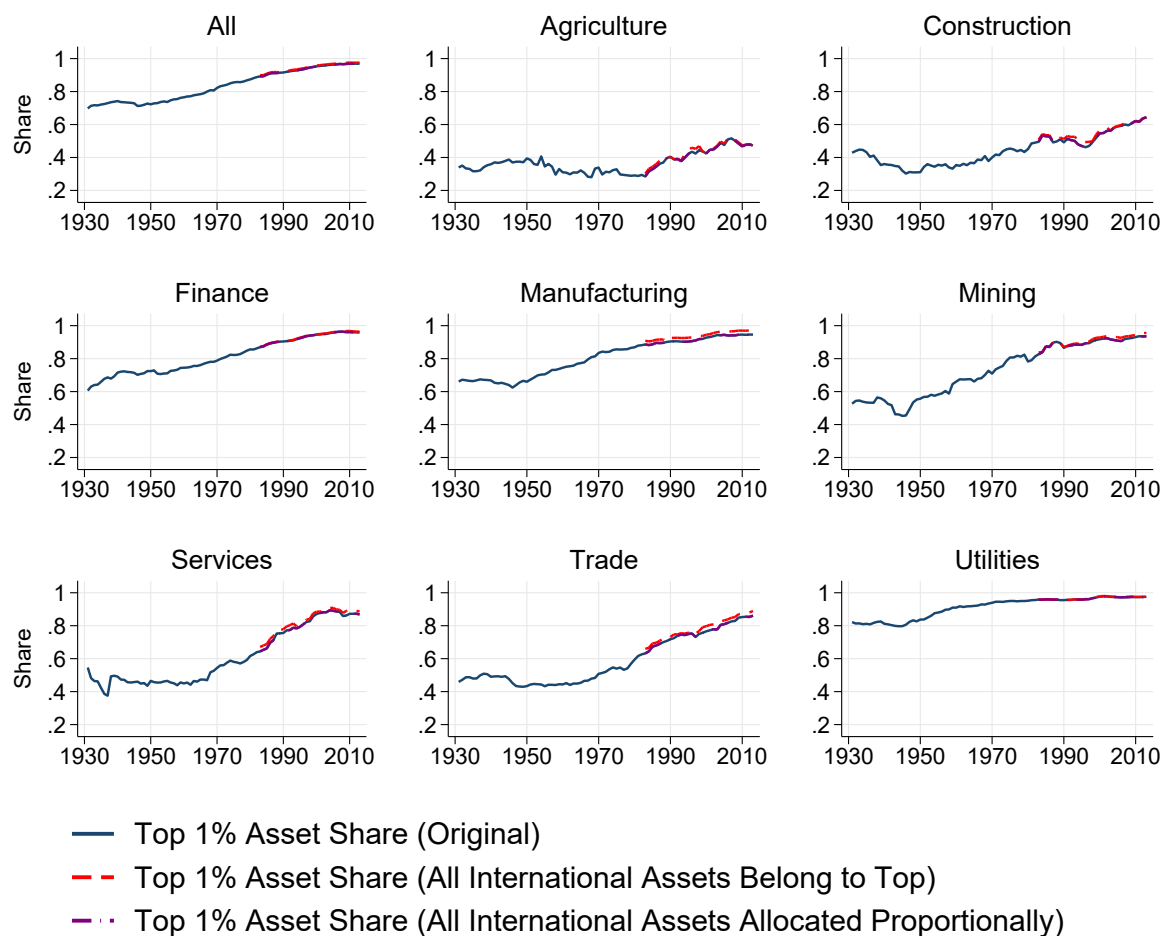


Figure IA9. Including International Assets

*Notes:* This figure shows estimated top 1% asset shares including international assets using Activities of U.S. Multinational Enterprises from the BEA. The solid line shows the original top 1% asset share using SOI data. The dashed line shows the top 1% asset share when all international assets are assigned to the top 1% businesses. The dash-dotted line shows the top 1% asset share when international assets are assigned to the top 1% and the bottom 99% businesses according to their domestic asset shares (using SOI data). See Internet Appendix [IA2.3](#) for details about variable construction.

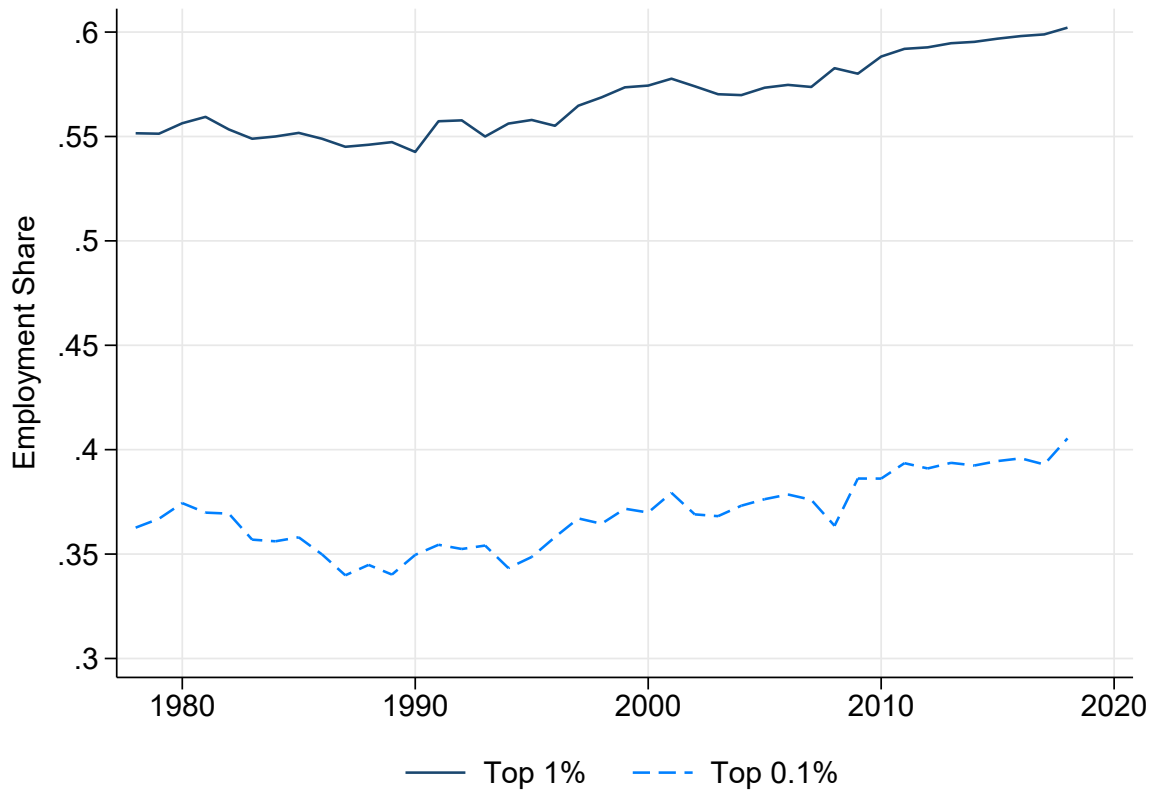


Figure IA10. Employment Concentration in Census Business Dynamic Statistics

*Notes:* This figure shows the aggregate employment share of the top 1% firms (solid line) and the top 0.1% firms (dashed line) by employment using census Business Dynamic Statistics. See Internet Appendix [IA2.3](#) for details about variable construction.

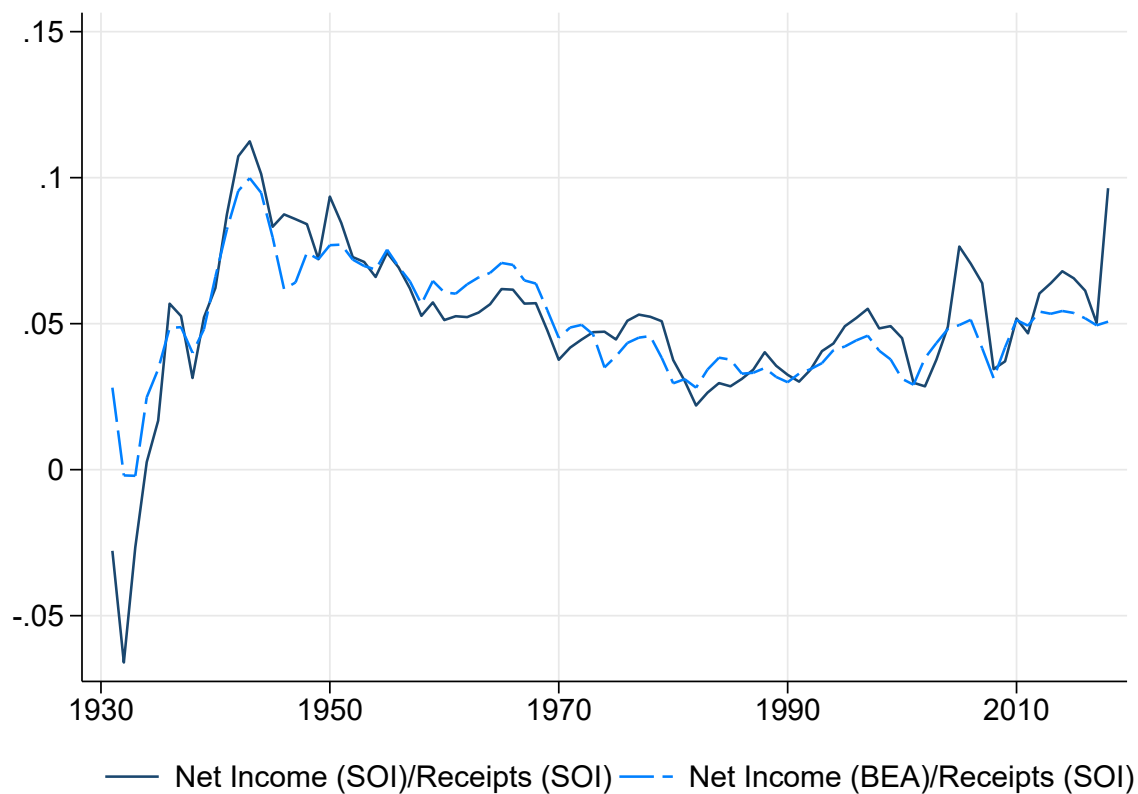


Figure IA11. Profitability in SOI and BEA

*Notes:* The solid line shows net income (before tax) in SOI normalized by total receipts in SOI. The dashed line shows net income (corporate profit before tax with inventory valuation and capital consumption adjustments) from BEA normalized by total receipts in SOI.



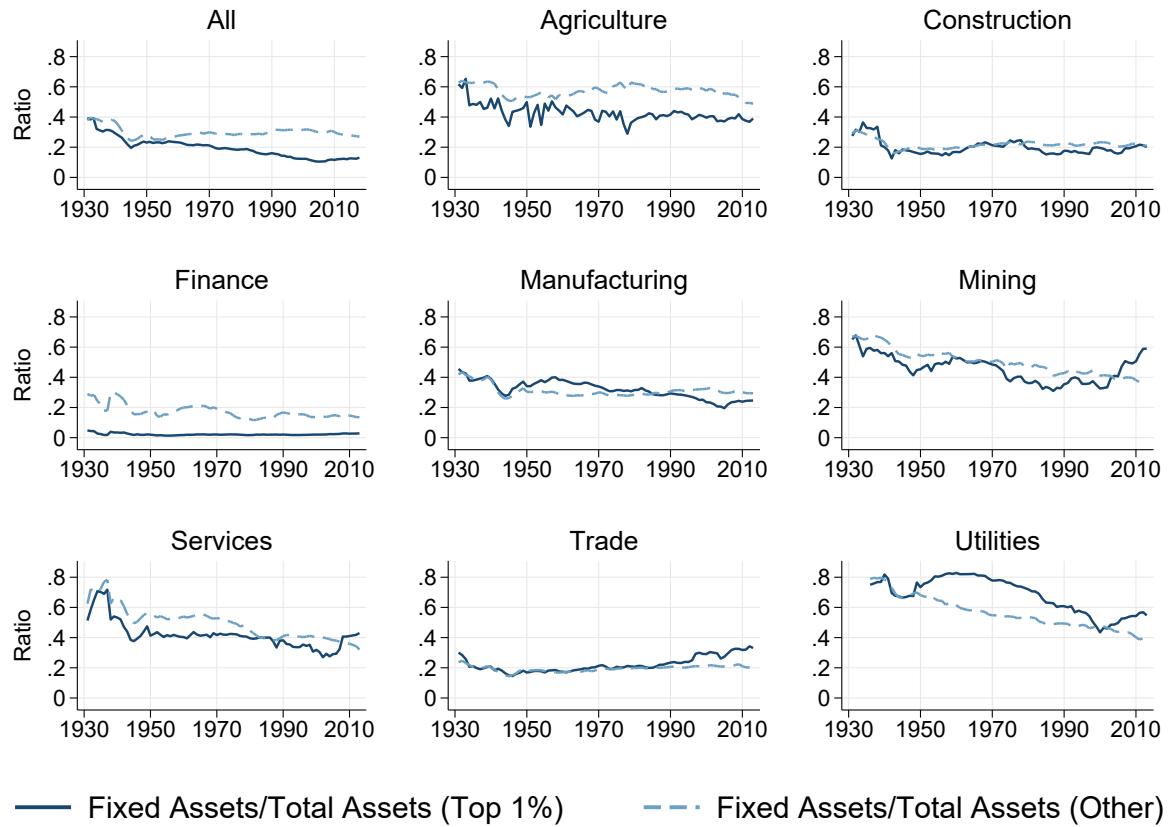


Figure IA12. Fixed Assets/Total Assets

*Notes:* This figure shows the ratio of fixed assets to total assets for the top 1% corporations by assets (solid line) and the rest (dashed line). Here we need to use the adding up bins method discussed in Section 2 to obtain fixed assets and total assets for these two groups (i.e., we add up fixed assets and total assets for each of these two groups). See Internet Appendix IA2.3 for details about variable construction.

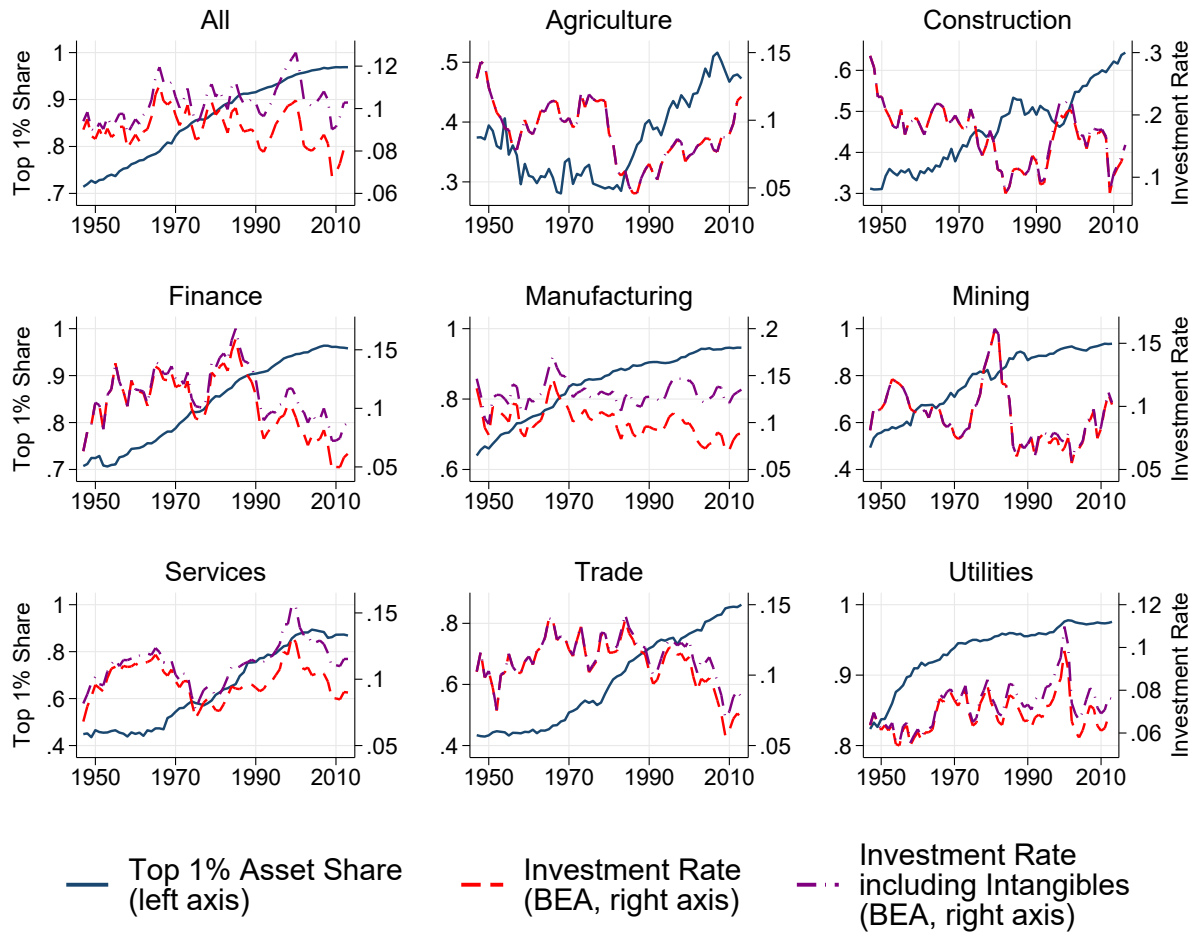


Figure IA13. Investment Rate

*Notes:* This figure shows the investment rate (investment over capital stock) from the BEA fixed asset tables. The dashed line shows the investment rate of fixed assets (equipment and structures). The dash-dotted line shows the investment rate of fixed assets plus intellectual property. The solid line is the asset share of the top 1% corporations by asset size from SOI. The left axis is the top 1% asset share, and the right axis is the investment rate. The mapping with industries in BEA fixed asset tables follows Table IA8.

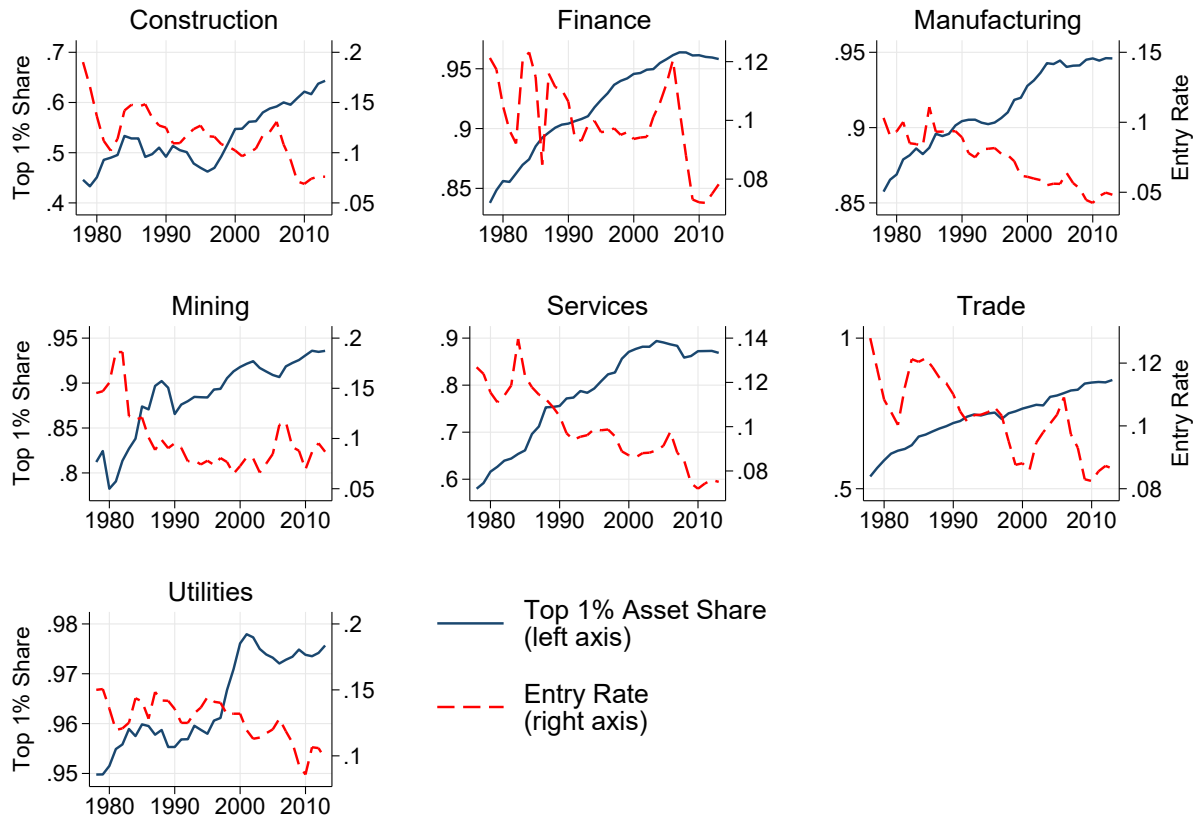


Figure IA14. Entry Rate

Notes: The dashed line shows the entry rate (the share of new firms) from census BDS. The solid line repeats the asset share of the top 1% corporations by asset size from SOI. The left axis is the top 1% asset share, and the right axis is the industry entry rate. The mapping with industries in BDS follows Table IA6.

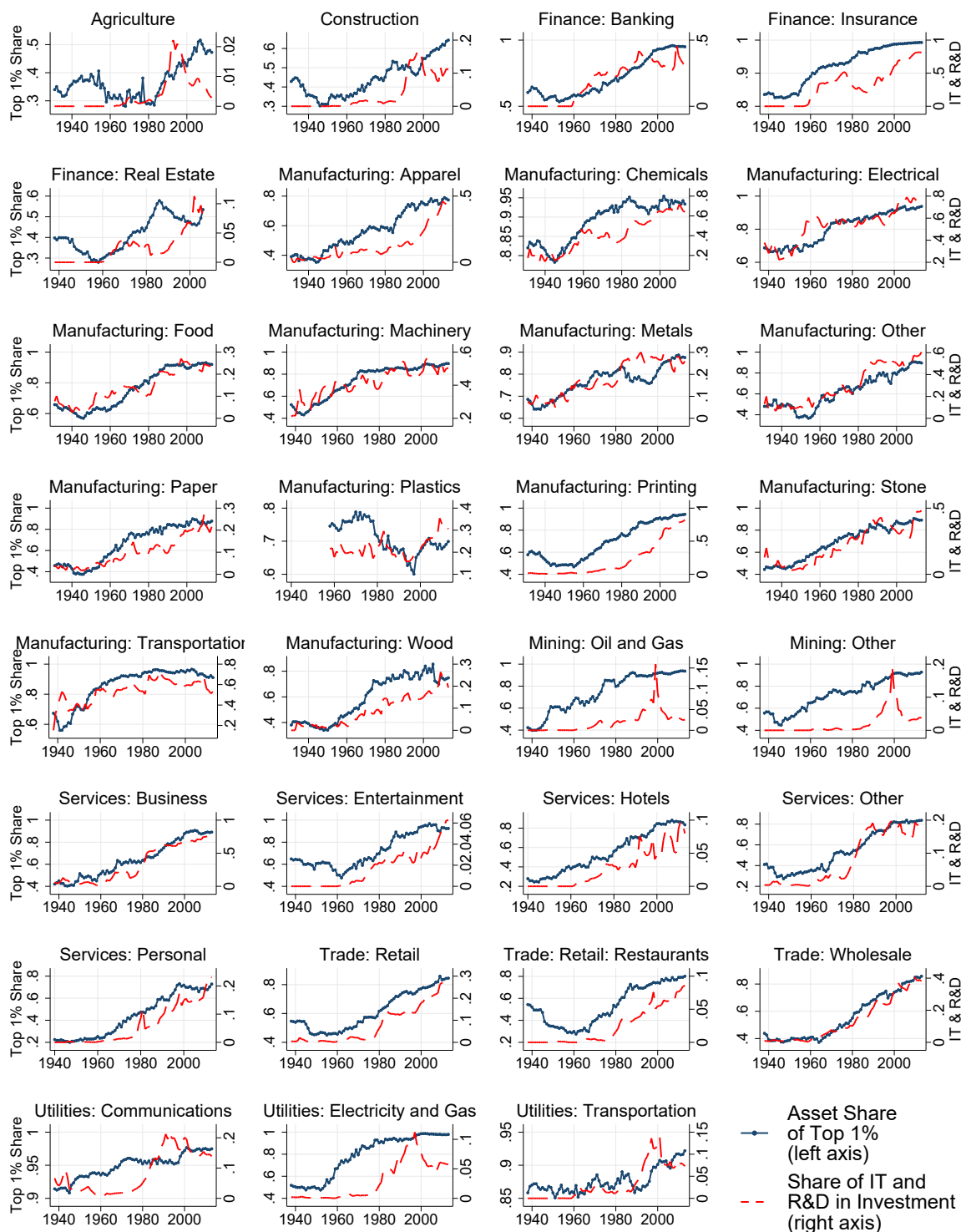


Figure IA15. Concentration and Technological Intensity

*Notes:* This figure shows the top 1% asset share (solid blue line) and the investment intensity of IT and R&D using BEA data (dashed red line). The left axis is the top 1% asset share and the right axis is the investment share in IT and R&D. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in BEA fixed asset tables follows Table IA8. The BEA data contain some noise. For instance, the value is sometimes zero in the early years. The BEA data also show large swings in a few industries, which could arise from changes in underlying data sources according to BEA staff.

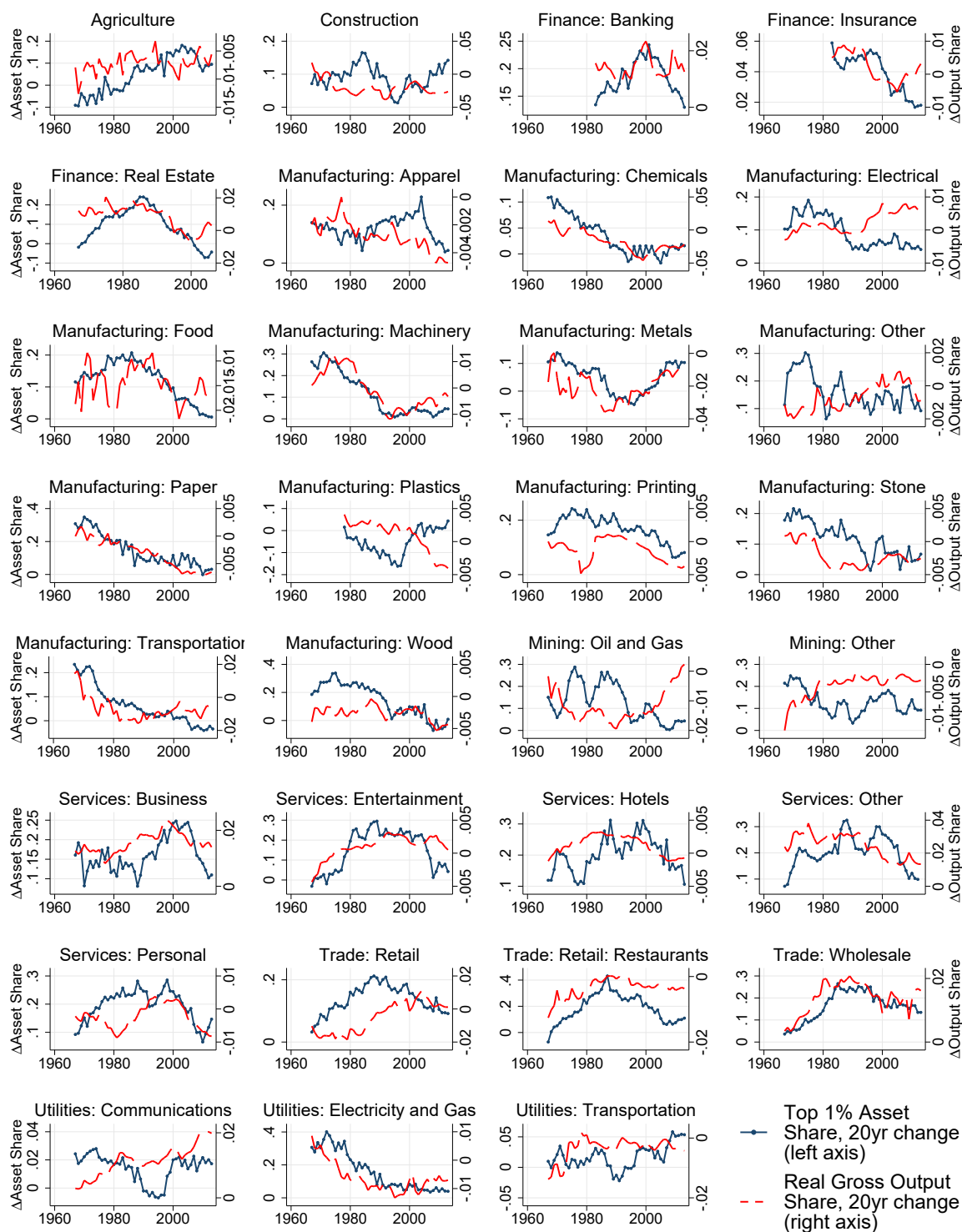


Figure IA16. Concentration and Industry Output Share

*Notes:* This figure shows changes in the top 1% asset share over twenty years (solid blue line) and changes in the industry's share in total real gross output (dashed red line). The left axis is changes in top 1% asset share, and the right axis is changes in the industry's share in total real gross output. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in the national accounts follows Tables [IA9](#) and [IA10](#).

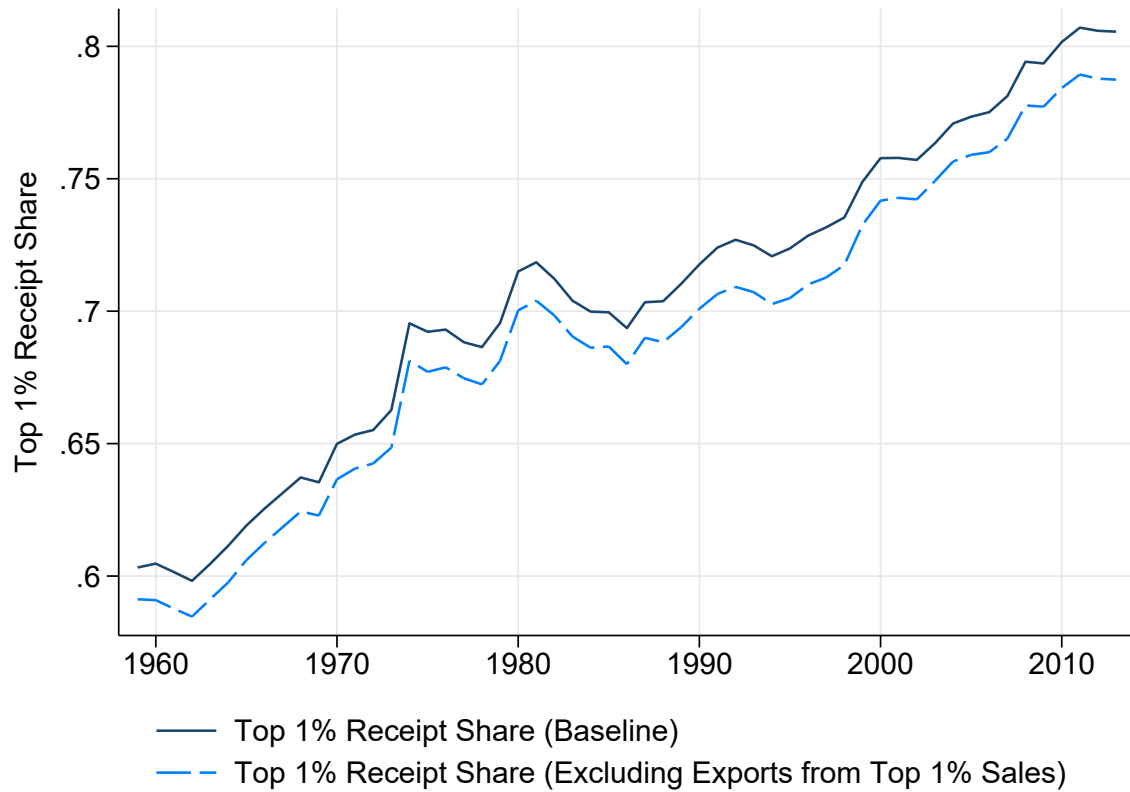


Figure IA17. Sales Concentration excluding Exports

*Notes:* The solid line shows the baseline aggregate sales concentration, which measures the share of the top 1% businesses by receipts in total receipts. The dashed line shows aggregate sales concentration excluding exports, where we remove all exports from the receipts of the top 1% businesses to be conservative. In other words, we calculate  $(\text{top 1\% receipts} - \text{exports}) / (\text{total receipts} - \text{exports})$ . See Internet Appendix [IA2.3](#) for details about variable construction.

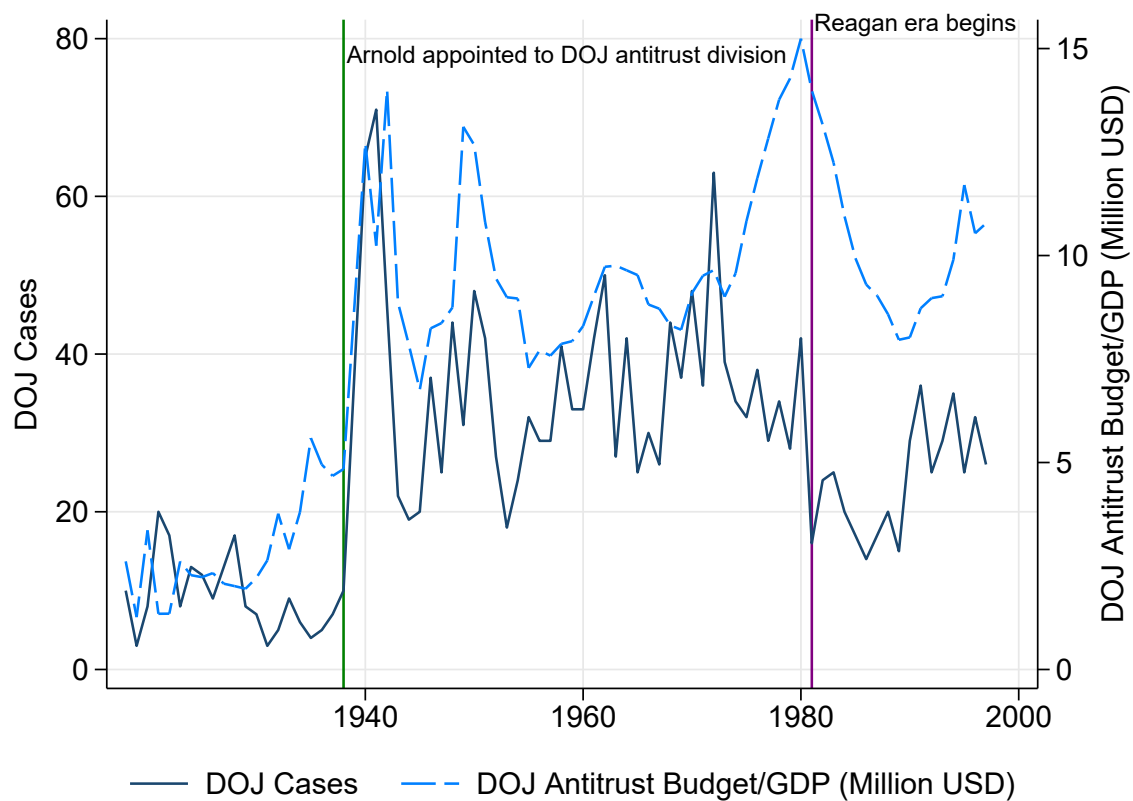


Figure IA18. DOJ Antitrust Activities

*Notes:* The solid line shows the annual number of antitrust cases institute by the DOJ. The dashed line shows the DOJ antitrust division budget normalized by GDP in million dollars.



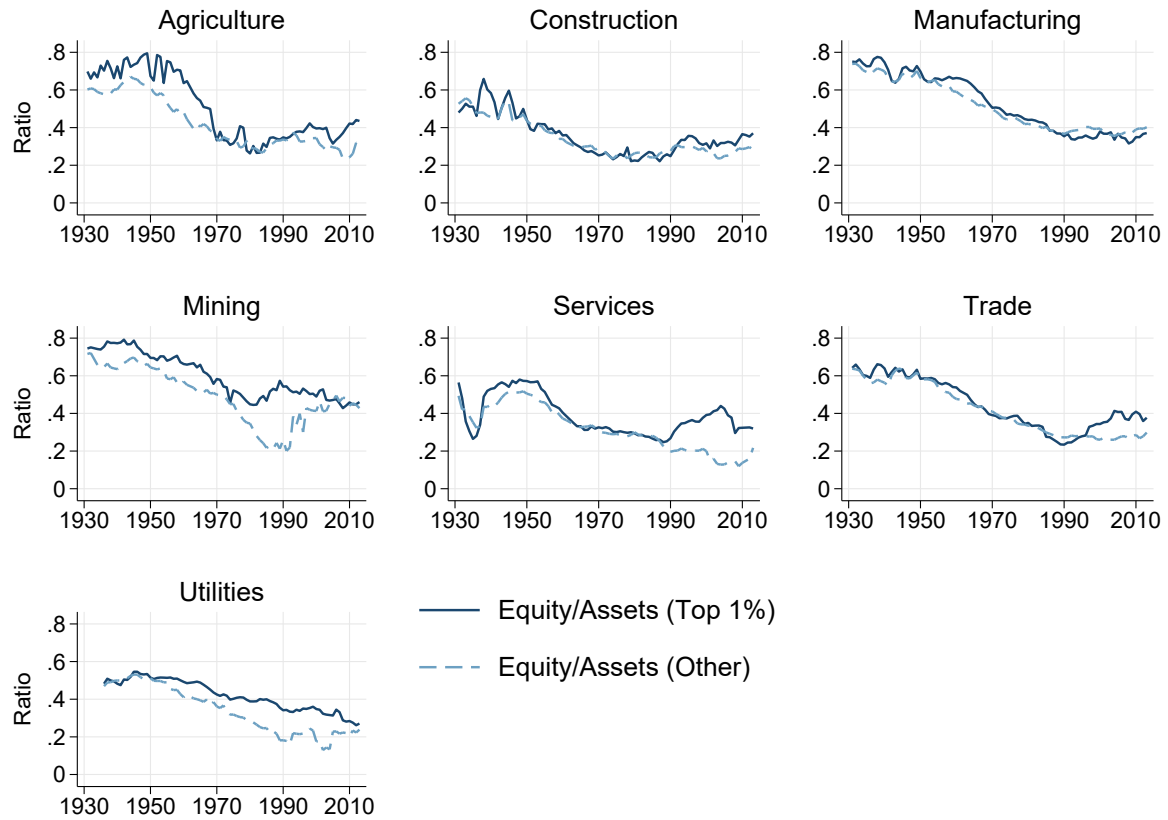


Figure IA19. Equity/Total Assets

*Notes:* This figure shows the ratio of book equity to total assets for the top 1% corporations by assets (solid line) and the rest (dashed line). Here we need to use the adding up bins method discussed in Section 2 to obtain book equity and total assets for these two groups (i.e., we add up book equity and total assets for each of these two groups). See Internet Appendix [IA2.3](#) for details about variable construction.

Table IA1 – Investment in IT and R&amp;D Normalized by Sales

|                                           | Top 1% Asset Share  |                     |                          |                     |                     |                     |                          |                     |
|-------------------------------------------|---------------------|---------------------|--------------------------|---------------------|---------------------|---------------------|--------------------------|---------------------|
|                                           | Level               |                     | Change ( $\Delta_{20}$ ) |                     | Level               |                     | Change ( $\Delta_{20}$ ) |                     |
|                                           | (1)                 | (2)                 | (3)                      | (4)                 | (5)                 | (6)                 | (7)                      | (8)                 |
| IT and R&D Investment/Sales               | 4.679***<br>(1.225) | 2.194***<br>(0.774) |                          |                     | 4.683***<br>(1.246) | 2.314***<br>(0.803) |                          |                     |
| $\Delta_{20}$ IT and R&D Investment/Sales |                     |                     | 0.964**<br>(0.430)       | 0.731***<br>(0.226) |                     |                     | 0.956**<br>(0.392)       | 0.777***<br>(0.218) |
| Industries                                | All                 |                     |                          |                     | Nonfinancial        |                     |                          |                     |
| Year Fixed Effect                         | No                  | Yes                 | No                       | Yes                 | No                  | Yes                 | No                       | Yes                 |
| Obs                                       | 2,394               | 2,394               | 1,774                    | 1,774               | 2,173               | 2,173               | 1,613                    | 1,613               |
| R <sup>2</sup>                            | 0.18                | 0.05                | 0.02                     | 0.01                | 0.20                | 0.07                | 0.02                     | 0.01                |

*Notes:* This table shows industry-level regressions of the top 1% asset share on the investment intensity of IT and R&D. The intensity of IT and R&D is measured using the annual spending on IT (computer equipment and software) and R&D in BEA fixed asset tables, normalized by total annual business receipts in SOI data. For both left hand side and right hand side variables, we use their levels in columns (1), (2), (5), and (6), and their changes over twenty years in columns (3), (4), (7), and (8). Year fixed effects are included in the even columns. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in BEA fixed asset tables follows Table IA8. Standard errors are Driscoll and Kraay (1998) with twenty lags.  $R^2$  does not include fixed effects. See Internet Appendix IA2.3 for details about variable construction.

Table IA2 – Robustness Check Controlling for the Number of Corporations

|                                       | $\Delta_{20}$ Top 1% Asset Share |                     |                    |                     |                     |                     |                     |                     |
|---------------------------------------|----------------------------------|---------------------|--------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                                       | (1)                              | (2)                 | (3)                | (4)                 | (5)                 | (6)                 | (7)                 | (8)                 |
| $\Delta_{20}$ Intensity of IT and R&D | 0.123**<br>(0.059)               | 0.083***<br>(0.031) | 0.123**<br>(0.059) | 0.083***<br>(0.031) |                     |                     |                     |                     |
| $\Delta_{20}$ Log Real Gross Output   |                                  |                     |                    |                     | 0.060***<br>(0.010) | 0.052***<br>(0.013) | 0.060***<br>(0.010) | 0.052***<br>(0.013) |
| $\Delta_{20}$ Log # of Corporations   | 0.006<br>(0.022)                 | -0.008<br>(0.027)   | 0.006<br>(0.022)   | -0.008<br>(0.027)   | 0.011<br>(0.020)    | -0.002<br>(0.021)   | 0.011<br>(0.020)    | -0.002<br>(0.021)   |
| Industries                            | All                              |                     | Nonfinancial       |                     | All                 |                     | Nonfinancial        |                     |
| Year Fixed Effect                     | No                               | Yes                 | No                 | Yes                 | No                  | Yes                 | No                  | Yes                 |
| Obs                                   | 1,774                            | 1,774               | 1,774              | 1,774               | 1,406               | 1,406               | 1,406               | 1,406               |
| R <sup>2</sup>                        | 0.02                             | 0.01                | 0.02               | 0.01                | 0.07                | 0.03                | 0.07                | 0.03                |

*Notes:* This table shows industry-level regressions of the twenty-year change of the top 1% asset share on the investment intensity of IT and R&D in columns (1) to (4), which are the same as columns (3), (4), (7) and (8) in Table 4 except we also control for the log change in the number of corporations in the industry. The table shows industry-level regressions of the twenty-year change of the top 1% asset share on industry growth over twenty years in columns (5) to (8), which are the same as columns (1), (2), (4), and (5) in Table 6 except we also control for the log change in the number of corporations in the industry. Year fixed effects are included in the even columns. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in BEA fixed asset tables follows Table IA8. Standard errors are Driscoll and Kraay (1998) with twenty lags.  $R^2$  does not include fixed effects.

Table IA3 – Technological Innovations Measured by Breakthrough Patents

|                                        | $\Delta_{20}$ Top 1% Asset Share |                  |                     |                      |                    |                     |                     |                      |
|----------------------------------------|----------------------------------|------------------|---------------------|----------------------|--------------------|---------------------|---------------------|----------------------|
|                                        | (1)                              | (2)              | (3)                 | (4)                  | (5)                | (6)                 | (7)                 | (8)                  |
| $\Delta_{20}$ Log Breakthrough Patents | 0.012<br>(0.011)                 | 0.007<br>(0.008) | 0.022***<br>(0.006) | 0.027***<br>(0.006)  | 0.028**<br>(0.011) | 0.016***<br>(0.004) | 0.018***<br>(0.004) | 0.021***<br>(0.004)  |
| $\Delta_{20}$ Log # of Patents         |                                  | 0.032<br>(0.025) |                     | -0.043***<br>(0.011) |                    | 0.063**<br>(0.028)  |                     | -0.043***<br>(0.013) |
| Sample Period                          | Full                             |                  |                     |                      | Pre-1980           |                     |                     |                      |
| Year Fixed Effect                      | No                               | No               | Yes                 | Yes                  | No                 | No                  | Yes                 | Yes                  |
| Obs                                    | 867                              | 867              | 867                 | 867                  | 453                | 453                 | 453                 | 453                  |
| R <sup>2</sup>                         | 0.02                             | 0.03             | 0.04                | 0.06                 | 0.06               | 0.12                | 0.03                | 0.04                 |

*Notes:* This table shows industry-level regressions of changes in the top 1% asset share on log changes in breakthrough patents in the industry (over twenty years). Breakthrough patent data come from [Kelly et al. \(2021\)](#) and the number of breakthrough patents is normalized by population. The even columns control for log changes in the number of patents in the industry. Year fixed effects are included in columns (3), (4), (7), and (8). Industries include manufacturing subsectors, mining subsectors, agriculture, construction, and utilities, due to the coverage of the breakthrough patent data. The mapping with industries in the patent data follows Table [IA6](#). Standard errors are [Driscoll and Kraay \(1998\)](#) with twenty lags.  $R^2$  does not include fixed effects.

Table IA4 – Robustness Check for Rising Concentration and Industry Growth

|                                     | $\Delta_{10}$ Top 1% Asset Share |                    |                    |                    |
|-------------------------------------|----------------------------------|--------------------|--------------------|--------------------|
|                                     | (1)                              | (2)                | (3)                | (4)                |
| $\Delta_{10}$ Log Real Gross Output | 0.122**<br>(0.048)               | 0.103**<br>(0.042) | 0.117**<br>(0.051) | 0.096**<br>(0.042) |
| Industries                          | All                              |                    | Nonfinancial       |                    |
| Year Fixed Effect                   | No                               | Yes                | No                 | Yes                |
| Obs                                 | 1,406                            | 1,406              | 1,304              | 1,304              |
| R <sup>2</sup>                      | 0.03                             |                    | 0.04               |                    |

*Notes:* This table shows industry-level regressions of changes in the top 1% asset share over ten years on industry real gross output growth over ten years predicted by the growth over the past ten years. Columns (1) and (2) show results for all industries. Columns (3) and (4) show results for nonfinancial industries. Year fixed effects are included in the even columns. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). The mapping with industries in the national accounts follows Tables [IA9](#) and [IA10](#). Standard errors are [Driscoll and Kraay \(1998\)](#) with twenty lags.  $R^2$  does not include fixed effects.

Table IA5 – Antitrust Enforcement

## Panel A. DOJ Antitrust Enforcement and Political Control over Presidential Cycles

|                      | Annual Cases        |                      | Annual Budget     |                     |
|----------------------|---------------------|----------------------|-------------------|---------------------|
|                      | (1)                 | (2)                  | (3)               | (4)                 |
| Republican President | -2.442<br>(4.997)   |                      | -0.224<br>(1.146) |                     |
| Republican Congress  | -4.069**<br>(1.878) |                      | -0.660<br>(0.464) |                     |
| Republican Trifecta  |                     | -4.878***<br>(1.877) |                   | -0.938**<br>(0.380) |
| Obs                  | 20                  | 20                   | 28                | 28                  |
| R <sup>2</sup>       | 0.22                | 0.25                 | 0.11              | 0.19                |

## Panel B. Changes in Concentration over Presidential Cycles

|                             | Change in Top 1% Asset Share |                    |                     |                     |
|-----------------------------|------------------------------|--------------------|---------------------|---------------------|
|                             | (1)                          | (2)                | (3)                 | (4)                 |
| Annual DOJ Cases            | -0.0002<br>(0.0006)          |                    |                     |                     |
| Annual DOJ Antitrust Budget |                              | 0.0004<br>(0.0016) |                     |                     |
| Republican President        |                              |                    | 0.0099<br>(0.0069)  |                     |
| Republican Congress         |                              |                    | -0.0000<br>(0.0001) |                     |
| Republican Trifecta         |                              |                    |                     | -0.0001<br>(0.0001) |
| Obs                         | 480                          | 571                | 571                 | 571                 |
| R <sup>2</sup>              | 0.00                         | 0.00               | 0.02                | 0.00                |

*Notes:* Panel A presents time series regressions of DOJ antitrust enforcement activities on political control. The outcome variables are the average annual DOJ antitrust cases (Annual Cases) and the average annual DOJ antitrust division budget normalized by GDP (Annual Budget) in each presidential cycle. The independent variables include an indicator for Republican presidential control (Republican President), the number of years with Republican control of the House as well as the Senate (Republican Congress), and the number of years with Reduplication control of both the presidency and congress (Republican Trifecta). Standard errors are Newey-West. Panel B presents panel regressions of changes in concentration in each industry on DOJ antitrust enforcement activities and political control. The outcome variable is the change in the top 1% asset share in each industry during each presidential cycle. The independent variables include the average annual DOJ antitrust cases and the average annual DOJ antitrust division budget (normalized by GDP) in each presidential cycle. The independent variables also include variables for Republican control. Each industry is a subsector, plus agriculture and construction (which do not have subsectors within them). Standard errors are [Driscoll and Kraay \(1998\)](#) with five lags.

## IA2 Data Construction

### IA2.1 SOI Data

We digitize data from historical publications of the Internal Revenue Service (IRS). The IRS has a longstanding tradition of collecting detailed statistics for individuals and businesses going back to the Revenue Act of 1916. The Statistics of Income (SOI) was first published in 1918 (with data for 1916). Initially the SOI included only basic statistics on corporations, but over the years the section on corporations has become increasingly detailed, with more cross-tabulations and variables. In addition to data on receipts and net income, the SOI also contains data on balance sheets, which derives from (end-of-fiscal-year) balance sheets submitted by corporations with their tax returns. Using micro data from these submissions, the SOI provides tabulations of businesses by size of net income and sector since 1918 (which ended in the 1970s), by size of assets and sector since 1931, and by business receipts and sector since 1959. We use these size tabulations to study trends in corporate concentration over the long run. As discussed in Section 2, the tabulations by size are mainly available for corporations (both C-corporations and S-corporations), and we provide additional checks for concentration estimates including noncorporations in Section 3.1 when we can obtain relevant data.

The SOI publications are accompanied by the Corporation Source Book, which is a series of initially unpublished volumes containing tabulations with more detailed classifications compared to the published reports. The Corporation Source Book is digitally available through the IRS and the Electronic Records Division at the U.S. National Archives and Records Administration from 1964. The advantage of the Corporation Source Book data is that it includes more granular sector data and additional income and balance sheet items. We use the Corporation Source Book whenever available.

**Scope** The scope of the SOI business size tabulations is all active corporations organized for profit that are required to file one of the 1120 forms. Different from the economic census, the SOI does not exclude specific industries (such as agriculture and rail transportation).<sup>29</sup> In contrast to individuals (Piketty and Saez, 2003), corporations have been required to file tax returns regardless of their income throughout our sample period. The earliest SOI publications were based on the analysis of all submitted corporate tax returns. In later years, the SOI used estimates from sample data. Starting in 1951, the IRS began to use a stratified probability sample to provide estimates for the whole population. In these samples the IRS varied the sampling rate by size (measured using the size of total assets or the size of net income at the beginning and the size of total assets or the size of business receipts more recently) to guarantee reliable totals. The sample usually included the universe of businesses in the top bins. Therefore, the transition to sample data should not affect our measurement of corporate concentration.

A small fraction of companies do not submit information about their balance sheets together with their tax returns. Reports without balance sheets are usually from corporations without assets (liquidations, dissolutions, acquisitions), foreign corporations doing business in the United States, and a small number of corporations that fail to supply balance sheet information. Until the SOI of 1958-59, these filings are included in all tabulations by net income, but excluded from tables pertaining to balance sheet information. Starting in 1959-60, the IRS included businesses with zero assets in the balance sheet tabulations and imputed data for businesses with missing balance sheets using information from

---

<sup>29</sup>The industries covered by the economic census can be seen here: <https://www.census.gov/programs-surveys/economic-census/technical-documentation/methodology.html#survey-design>.

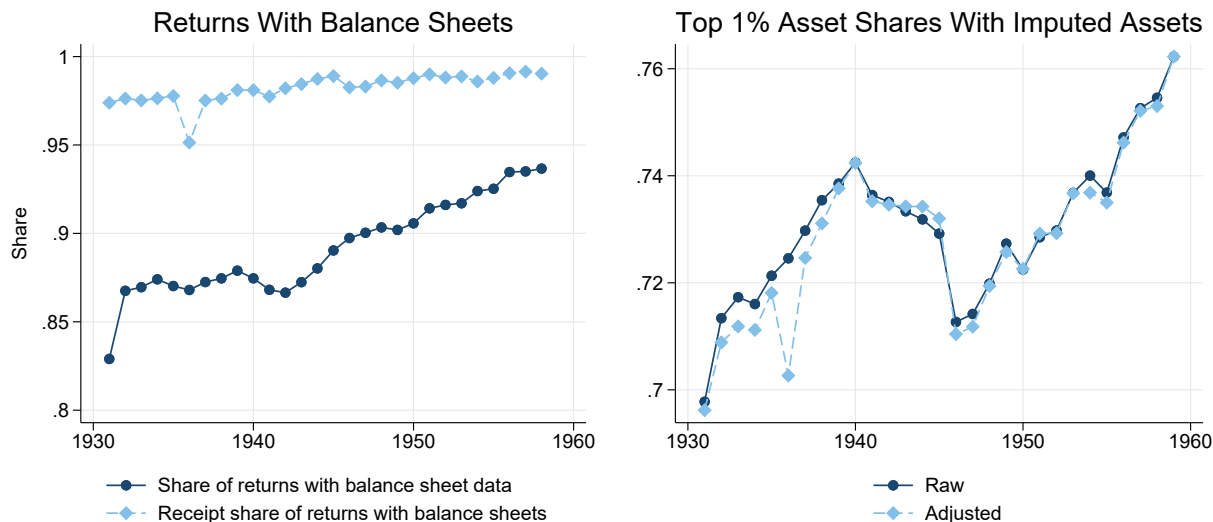


Figure IA20. Returns without Balance Sheets

*Notes:* The right panel shows share of returns submitted with balance sheet information and the receipt share accounted for by these returns. The left panel shows how concentration changes if we impute assets for firms without balance sheet information using information on their receipts and assuming that they have the same assets-to-receipts ratios as the industry as a whole.

the returns of businesses with both income statements and balance sheets in the same industry. Taken together, before 1959, the omission of businesses with missing balance sheet information in the SOI asset bin tabulations could affect the number of businesses in our calculations (for the asset share of top businesses). The left panel of Figure IA20 shows the share of returns in each year with balance sheet information and the receipt share accounted for by these returns. For example, in 1950 about 10 percent of returns representing 1.2 percent of total receipts did not include balance sheet data. The figure also shows that both the share of returns without balance sheet data and their receipt share has declined over time. We can provide robustness checks by either assuming that the businesses with missing balance sheets fall in the smallest asset size bin, or imputing the asset size bins they belong to using information on their receipts (assuming they have the same assets-to-receipts ratios as the industry as a whole). The right panel of Figure IA20 compares our baseline concentration estimate to a concentration estimate with imputed assets for returns without balance sheets. The two series are similar. We also find the same degree of consistency at the industry level (results not shown).

**Industry classification** The SOI assigns a single industry code to each business based on the industry that represents the largest percentage of its total business receipts. For studies using long-run data by industry, a common task is to address changes to the industry classification systems over time. We harmonize the different industry classification systems to construct consistent industries. The SOI industry classification can be broadly separated into three periods. Between 1931 and 1937, the IRS followed its own industry classification. In 1938, the IRS adopted the newly created SIC industry classification system (with a few small modifications), and followed its various vintages until 1997. In 1998, the IRS began to use NAICS codes. Broad industrial groupings remained relatively stable within these three periods, which allows us to build consistent definitions for main sectors (roughly at the level of one-digit SIC codes) and subsectors (roughly at the level of two-digit SIC codes).



Table [IA6](#), Panel A, presents how our main sectors correspond to Industrial Divisions in the SIC classification system and NAICS codes. Panel B shows the construction of the subsectors. These subsectors are also designed to maximize the comparability with industries in BEA data (including the BEA fixed asset tables and NIPA accounts), since our analyses also rely on BEA data to measure various outcomes. If we are not mapping into industries in BEA data, then we can further break down several subsectors. Among “Construction,” we can have “Construction: Buildings” (SIC 15, NAICS 236), “Construction: Heavy Construction” (SIC 16, NAICS 237), and “Construction: Special Trade” (SIC 17, NAICS 238). Among Mining, we can have “Mining: Metal” (SIC 10, NAICS 2122), “Mining: Coal” (SIC 12, NAICS 2121), and “Mining: Non Metallic” (SIC 14, NAICS 2123). Among “Manufacturing: Apparel,” we can have “Manufacturing: Apparel and Textiles” (SIC 22 and 23, NAICS 313, 314, and 315) and “Manufacturing: Leather” (SIC 31 and NAICS 316). Among “Trade: Retail,” we can have “Trade: Retail: Apparel” (SIC 56, NAICS 448), “Trade: Retail: Automotive” (SIC 55, NAICS 441 and 447), “Trade: Retail: Building Materials” (SIC 52, NAICS 444), “Trade: Retail: Food” (SIC 54, NAICS 445), “Trade: Retail: Furniture” (SIC 57, NAICS 442), “Trade: Retail: General Merchandise” (SIC 53, NAICS 452) and “Trade: Retail: Miscellaneous” (SIC 59, NAICS 446, 451, 453, and 454). Among “Services: Other,” we can have “Services: Repair” (SIC 75 and 76, NAICS 532 and 811) and “Services: Miscellaneous” (SIC 89, NAICS 561, 61, 62, and 813). Finally, because we have more granular tabulations for businesses by asset size than businesses by receipt size, for the former we can make more refined mapping when the SOI data transitioned from the SIC classification system to NAICS codes.

**Bin deletion** For certain size bins at the industry level, financial data are suppressed to avoid disclosing information of individual businesses. This problem rarely arises in the main sector data, but becomes more common at the subsector level. For some of the early SOI issues, we can manually back out the missing values using adding up constraints from the hierarchical industry and bin structure (similar in spirit to [Eckert et al., 2020](#)). In later years, additional precautions have been introduced by the IRS to preserve taxpayer confidentiality by deleting information from additional size and industry bins when necessary. In these cases, we join the deleted bins (and all bins in between) into one large bin, and back out the financial data using the difference of the total and all other bins. While this approach generally works well and does not create problems for the concentration measures, in a handful of cases the number of size bins is reduced too much to obtain consistent and robust top shares. We linearly interpolate data for these years.

**Consolidation** The IRS allows corporations to file consolidated returns if at least 80 percent of the equity of an affiliate is owned within the group. Corporations that chose to file consolidated returns in one year are generally also required to file consolidated returns in the subsequent years. The consolidation privilege is granted to all affiliated domestic corporations except regulated investment companies (RICs), real estate investment trusts (REITs), tax-exempt corporations, Interest Charge Domestic International Sales Corporations (IC-DISCs), and S-corporations. Life insurance companies can file consolidated returns with other life insurance companies without restrictions. In recent years at least, eligible firms generally elect to consolidate ([Mills, Newberry, and Trautman, 2002](#)), given more favorable treatments when consolidated (e.g., when consolidated the sales among affiliates do not generate taxes, and gains and losses across affiliates can be netted).

Rules on consolidation for tax purposes have had several changes over time. [Streuling \(1971\)](#) offers a detailed discussion of the various Revenue Acts that led to the changes. First, the 80% ownership requirement applicable today dates back to 1954. Prior to 1954, the ownership threshold was 95%. Second, consolidated returns were often taxed at higher rates before the 1960s. In 1932 and 1933,

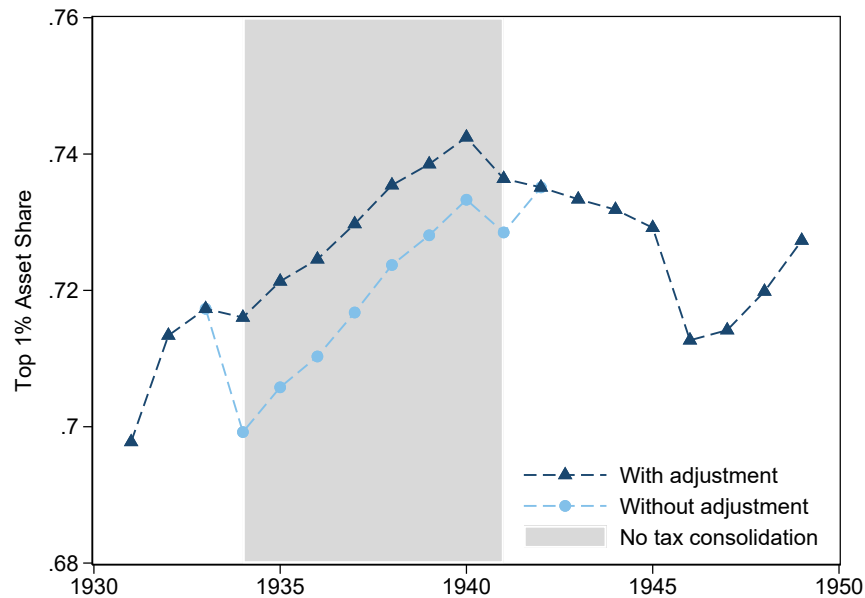


Figure IA21. Consolidation Adjustment

*Notes:* This figure shows the top 1% asset share between 1931 and 1950 with and without adjustment for changes in consolidation.

consolidated returns were subject to an additional tax of 0.75 percent. In 1934 and 1935, the additional tax increased to 1 percent. No additional tax was imposed between 1936 and 1941, but the consolidation privilege was significantly limited (see below). Between 1942 and 1963, corporations filing consolidated returns were subject to a surtax on the group of two percentage points. The Revenue Act of 1964 eventually repealed the two percent surtax for consolidated returns, so surtaxes no longer applied since 1964. Finally, consolidation was mandatory between 1918 and 1921 and voluntary after 1922. Then between 1934 and 1941, there was a change in procedure whereby all corporations (except for railway companies that were affiliated with each other) were not allowed to file consolidated returns. This change led to an upward shift in the number of returns and a downward shift in concentration. While this policy change only induced a relatively modest decline in the top 1% asset share for the whole economy (see Figure IA21), its effects in sectors with many consolidated returns (particularly Utilities and Manufacturing: Chemicals) were more sizable.

We adjust the 1934 to 1941 concentration estimates for all sectors using two approaches. First, if we have data before 1934 and after 1942, then we scale the 1934 to 1941 data using the 1933 and 1942 benchmarks and divide the remaining level difference equally over the 1934 to 1941 period. This allows us to rescale the data to the correct level, while preserving the time trends of the 1934 to 1941 period. Second, for some subsectors, our concentration estimates only begin in 1938 (with the introduction of SIC industry codes). For these sectors, we assume that concentration did not change between 1941 and 1942 and rescale earlier years accordingly. The effects of our adjustment can be seen in Figure IA21. The dashed line with circles shows the top 1% asset shares without adjustment and the dashed line with triangles shows the adjusted series.

One possible concern is that changes in the prevalence of consolidation may affect the concentration

trends we observe. For the early years (e.g., 1929 and 1933), [Means et al. \(1939\)](#) used SOI micro data to manually consolidate the unconsolidated subsidiaries of the largest 200 nonfinancial corporations. They then calculated the share of the top 200 nonfinancial corporations in total corporate assets (among other things) after the adjustment. Their results were similar to our calculation of the top 0.1% share among nonfinancial corporations for those years (during that time nonfinancial industries had slightly over 200 corporations for the top 0.1%). We make three observations for the subsequent years. First, we digitize data on the share of consolidated returns in total returns using information about consolidated returns in the SOI. Figure [IA22](#) shows the share of consolidated returns in the total number of returns (dark blue circles), and the share of assets from consolidated returns in total assets (light blue diamonds). We observe a decrease in the prevalence of consolidated returns between the early 1930s and the 1940s. Then the prevalence of consolidated returns increased from the mid-1960s to the 1980s, roughly returning to the prevalence of consolidated returns in the early 1930s. Meanwhile, top 1% asset shares were much higher in the 1980s relative to the 1930s. After the 1980s, the prevalence of consolidated returns decreased in number (though not much in their shares of total assets), while top 1% shares continued to rise.

Second, within each subperiod of consolidation rules (1934 to 1941, 1942 to 1954, 1954 to 1964, and after 1964), we generally observe rising top 1% asset shares, as shown in Figure [IA23](#). Here we present the final top 1% asset shares in our data, using manufacturing and aggregate series as examples. The only modification to the raw results from the SOI is the adjustment for the 1934 to 1941 period as explained above.

Finally, the consolidation rules apply to all sectors and the consolidation trends are largely similar across sectors, but the concentration trends display differences in the timing of rising concentration. In the analyses of the mechanisms behind rising concentration in Section [4.1](#), we use time fixed effects to isolate the timing differences in rising concentration across industries (see Tables [4](#) to [6](#)); these time fixed effects should absorb the impact of changes in consolidation rules which apply to all industries.

**Partnerships and sole proprietorships** We also collect data for noncorporations (partnerships and sole proprietorships) from historical SOI publications. We have information about the number of noncorporations and their total receipts (using the dataset compiled by [Lamoreaux \(2006\)](#), extended to recent years using data from Table 3 of the Integrated Business Statistics on the IRS website). In addition, we collect information on the total net income of noncorporations (restricted to noncorporations with positive net income) from historical SOI publications between 1945 and 1974. After 1957 the IRS began to publish annual SOI reports on noncorporations. Before 1957, SOI reported net income for sole proprietorships biannually and published reports on partnerships in certain years (1945, 1947, 1953). We use information on partnership and sole proprietorship income from individual tax returns to interpolate the years in between.

For 1957 to 1980 and 1998 to 2003, we are able to obtain tabulations of noncorporations by size bins of business receipts for the main sectors. Table [IA7](#) shows the detailed list of the sources.<sup>30</sup> The tabulations separately report the number of returns without business receipts data in some years. We

---

<sup>30</sup>We exclude the main sector “Finance” for 1998 to 2003 due to an inconsistency in the tabulations for partnerships: the column on “gross receipts” has data errors and the values presented there do not represent gross receipts.

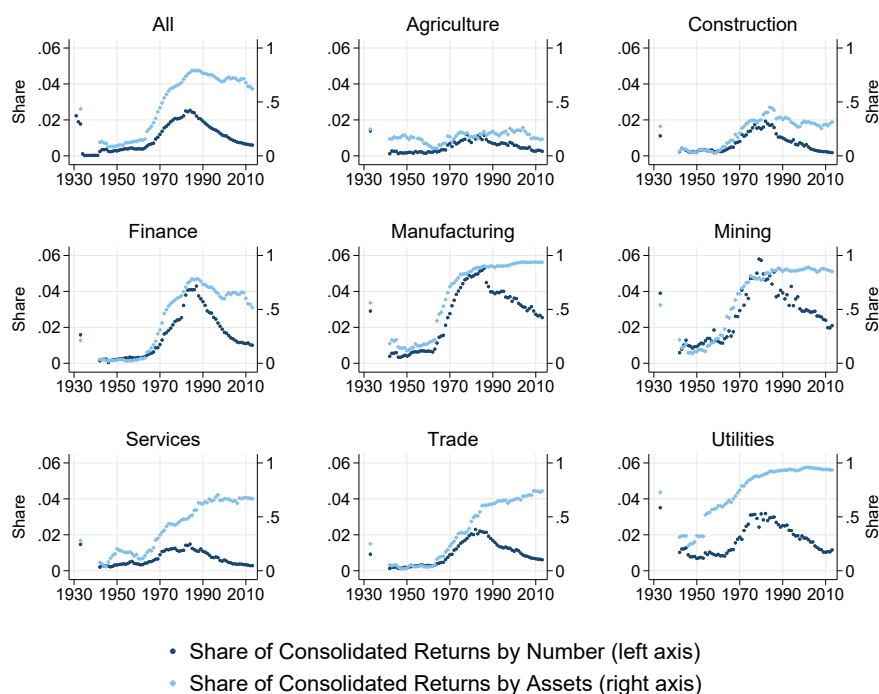


Figure IA22. Prevalence of Consolidation

*Notes:* This figure shows the prevalence of consolidation over time. The dark blue circles show the share of consolidated returns in the total number of returns, and the light blue diamonds show the share of assets from consolidated returns in total assets.

currently exclude these returns from our concentration estimates, but the results are not sensitive to this assumption (the concentration estimates are very similar if we assume the businesses without receipt data belong to the smallest bin, which appears to be the case).

Data on partnerships come from Form 1065, which partnerships need to file with the IRS for informational purposes. The partnership data include all groups conducting business for profit unless classified as corporations, trusts, or estates for tax purposes. These data include partnerships, syndicates, joint ventures and other unincorporated organizations. Data on sole proprietorships are based on the business schedule of owners' individual tax returns. Properly identifying the number of sole proprietorships from individual tax returns has always been a challenge for IRS statisticians. Since 1981, when a return has more than one business schedule, data from the schedules are combined to simplify statistical processing. This implies that the statistics effectively report the number and business receipts of proprietors, rather than those of individual proprietorships. However, the ratio between the number of companies and the number of owners is claimed to be relatively small (between 1 and 1.1, [Lamoreaux, 2006](#)). Prior to 1981, the SOI counting of sole proprietorships differed over the years, in particular when individuals filed multiple business schedules per individual return. In some years, the reporting unit was the number of C Schedules filed with the return, and in some other years the SOI only counted those businesses that operated in different industries as separate businesses, or restricted the overall number of businesses per owner.

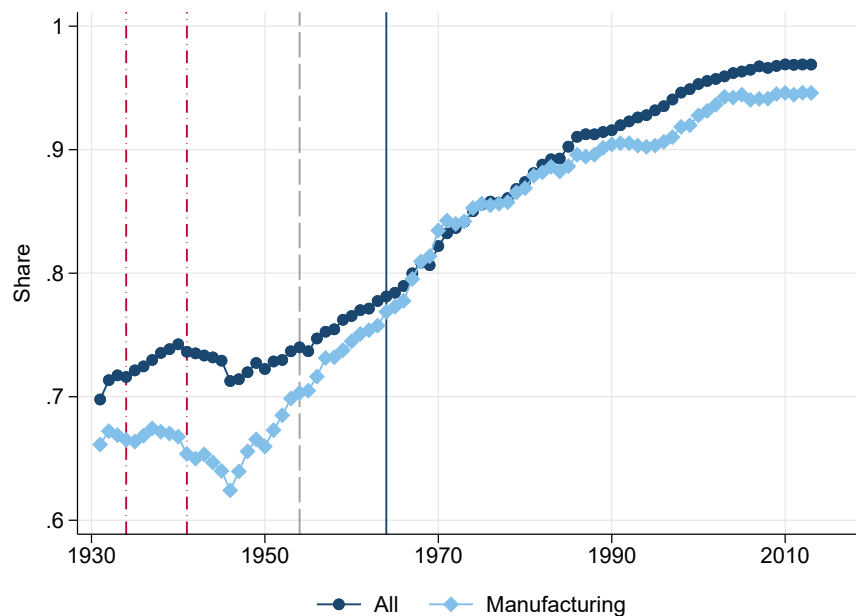


Figure IA23. Top 1% Asset Shares under Different Consolidation Rules

*Notes:* This figure shows the top 1% asset share for the aggregate economy (dark blue circles) and for manufacturing (light blue diamonds). The dash-dotted red lines mark the 1934 to 1941 period where consolidated filings were not allowed; the concentration estimates in this period use our adjustment explained above. The dashed gray line marks 1954, where the consolidation threshold changed from 95% ownership in affiliates to 80% ownership. The blue line marks 1964, where the surtax on consolidated returns ended.

Tax returns and schedules of partnerships, sole proprietorships, and corporations use different terms to describe items that are similar in nature. The historical SOI publications adopted common naming conventions across business organizations. For sole proprietorships, business receipts are defined as total receipts from sales and services, less rebates, returns, and allowances plus other business income. For partnerships, business receipts are defined as gross receipts from sales and services, less rebates, returns. For corporations, business receipts are defined as gross sales plus gross receipts from operations, less rebates, returns and allowances. For sole proprietorships, net income represents the difference between business receipts and the sum of cost of sales and operations and other business deductions. For partnerships, net income represents the difference between total taxable receipts and the sum of cost of sales and operations and other business deductions. For corporations, net income represents the difference between total receipts and total business deductions. Because reporting requirements differ across organizational forms, net income and receipts may not be strictly comparable. For instance, sole proprietorships are not allowed to deduct the owners' salary from their net income, and investment income is automatically treated as personal income instead of business income. For more detail on the noncorporation data, see the original documents on the IRS website (Table IA7) and the documentation of Lamoreaux (2006).

Table IA6 – Industry Harmonization of SOI Data

## Panel A. Main Sectors

| Main Sector   | SIC Industry Division                       | NAICS Codes                                                |
|---------------|---------------------------------------------|------------------------------------------------------------|
| Agriculture   | Agriculture, Forestry, Fishing (01-09)      | 11                                                         |
| Mining        | Mining (10-14)                              | 21                                                         |
| Construction  | Construction (15-17)                        | 23                                                         |
| Manufacturing | Manufacturing (20-39)                       | 31-33, 511                                                 |
| Utilities     | Transportation and Public Utilities (40-49) | 22, 48-49, 513, 515, 517, 562                              |
| Trade         | Wholesale and Retail Trade (50-59)          | 42-45, 722                                                 |
| Finance       | Finance, Insurance, and Real Estate (60-67) | 52, 531, 533, 55                                           |
| Services      | Services (70-89)                            | 512, 514, 516, 518, 519, 532, 54, 561, 61, 62, 71, 721, 81 |

## Panel B. Subsectors

| Subsector                      | SIC Industry Group                                                                                                                                      | NAICS Codes                |
|--------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|
| Finance: Banking               | Banking (60), Credit Agencies Other than Banks (61), Security and Commodity Brokers (62)                                                                | 522, 523                   |
| Finance: Holding Companies     | Holding and Other Investment Companies (67)                                                                                                             | 525, 55                    |
| Finance: Insurance             | Insurance (63)                                                                                                                                          | 524                        |
| Finance: Real Estate           | Real Estate (65)                                                                                                                                        | 531, 533                   |
| Manufacturing: Apparel         | Textile Mill Products (22), Apparel (23), Leather (31)                                                                                                  | 313, 314, 315, 316         |
| Manufacturing: Chemicals       | Chemicals and Allied Products (28)                                                                                                                      | 324, 325                   |
| Manufacturing: Electrical      | Electronic (36), Measuring, Analyzing, and Controlling Instruments (38)                                                                                 | 334, 335                   |
| Manufacturing: Food            | Food and Kindred Products (20), Tobacco Products (21)                                                                                                   | 311, 312                   |
| Manufacturing: Machinery       | Industrial and Commercial Machinery (35)                                                                                                                | 333                        |
| Manufacturing: Metals          | Primary Metal (33), Fabricated Metal Products (34)                                                                                                      | 331, 332                   |
| Manufacturing: Other           | Miscellaneous Manufacturing (39)                                                                                                                        | 339                        |
| Manufacturing: Paper           | Paper and Allied Products (26)                                                                                                                          | 322                        |
| Manufacturing: Plastics        | Rubber and Plastics Products (30)                                                                                                                       | 326                        |
| Manufacturing: Printing        | Printing, Publishing, and Allied Industries (27)                                                                                                        | 323                        |
| Manufacturing: Stone           | Stone, Clay, Glass, and Concrete Products (32)                                                                                                          | 327                        |
| Manufacturing: Transportation  | Transportation Equipment (37)                                                                                                                           | 336                        |
| Manufacturing: Wood            | Lumber and Wood Products (24), Furniture and Fixtures (25)                                                                                              | 321, 337                   |
| Mining: Oil and Gas            | Oil and Gas Extraction (13)                                                                                                                             | 211, 213                   |
| Mining: Other                  | Metal Mining (10), Coal and Lignite Mining (12), Nonmetallic Minerals (14)                                                                              | 212                        |
| Services: Business             | Business Services (73)                                                                                                                                  | 54, 514, 516, 518, 519     |
| Services: Entertainment        | Motion Pictures (78), Amusement and Recreation (79),                                                                                                    | 512, 71                    |
| Services: Hotels               | Hotels and Other Lodging Places (70)                                                                                                                    | 721                        |
| Services: Other                | Auto Repair (75), Miscellaneous Repair Services (76), Health Services (80), Legal Services (81), Educational Services (82), Miscellaneous Services (89) | 532, 561, 61, 62, 811, 813 |
| Services: Personal             | Personal Services (72)                                                                                                                                  | 812                        |
| Trade: Retail                  | Retail Trade (52-57, 59)                                                                                                                                | 44-45                      |
| Trade: Retail: Restaurants     | Eating and Drinking Places (58)                                                                                                                         | 722                        |
| Trade: Wholesale               | Wholesale Trade (50-51)                                                                                                                                 | 42                         |
| Utilities: Communications      | Communications (48)                                                                                                                                     | 513, 515, 517              |
| Utilities: Electricity and Gas | Electric and Gas (49)                                                                                                                                   | 22, 562                    |
| Utilities: Transportation      | Transportation (40-47)                                                                                                                                  | 48, 49                     |

Notes: This table shows the mapping between historical SOI industries and our main sectors and subsectors. SOI industries are classified by economic activity using SIC codes until 1997 and NAICS industry codes afterwards. The SOI sometimes departs from the SIC and NAICS classification systems in order to reflect particular provisions in the Internal Revenue Code. However, the SOI industries are generally very similar to SIC and NAICS industries, so we illustrate them using SIC codes (in the second column) and NAICS codes (in the third column). Panel A shows the main sectors in our data (the first column) and the correspondence with SIC industry divisions and NAICS industry codes. Panel B shows the subsectors in our data (the first column) and the correspondence with SIC industry groups and NAICS codes.

Table IA7 – Sources for Noncorporation Tabulations

| Year | Source                                                              | Business Type and Notes                |
|------|---------------------------------------------------------------------|----------------------------------------|
| 1957 | U.S. Business tax returns, July 1957-June 1958, pages 8ff. & 12ff   | All; only "All Industries"             |
| 1958 | U.S. Business tax returns, July 1958-June 1959, pages 22ff. & 36ff. | All                                    |
| 1959 | U.S. Business tax returns, July 1959-June 1960, pages 18ff. & 46ff. | All                                    |
| 1960 | U.S. Business tax returns, July 1960-June 1961, pages 24ff. & 56ff. | All                                    |
| 1961 | U.S. Business tax returns, July 1961-June 1962, pages 24ff. & 54ff. | All                                    |
| 1962 | U.S. Business tax returns, 1962, pages 34ff. & 120ff.               | All                                    |
| 1964 | U.S. Business tax returns, July 1964-June 1965, pages 87ff.         | Partnerships                           |
| 1965 | U.S. Business income tax returns, 1965, pages 39ff. & 129ff.        | All                                    |
| 1966 | U.S. Business income tax returns, 1966, pages 37ff. & 130ff.        | All                                    |
| 1967 | U.S. Business income tax returns, 1967, pages 41ff. & 165ff.        | All                                    |
| 1968 | U.S. Business income tax returns, 1968, pages 36ff. & 143ff.        | All                                    |
| 1969 | U.S. Business income tax returns, 1969, pages 34ff. & 115ff.        | All; IRS scan incomplete               |
| 1970 | U.S. Business income tax returns, 1970, pages 38ff. & 123ff.        | All                                    |
| 1971 | U.S. Business income tax returns, 1971, pages 35ff. & 123ff.        | All                                    |
| 1972 | U.S. Business income tax returns, 1972, pages 18ff. & 123ff.        | All                                    |
| 1973 | U.S. Business income tax returns, 1973, pages 23ff. & 144ff.        | All                                    |
| 1974 | U.S. Business income tax returns, 1974, pages 30ff. & 145ff.        | All                                    |
| 1975 | U.S. Business income tax returns, 1975, pages 22ff. & 220ff.        | All                                    |
| 1976 | U.S. Business income tax returns, 1976, pages 34ff. & 259ff.        | All                                    |
| 1977 | Sole Proprietorships returns, 1977, pages 42ff.                     | Sole proprietorships                   |
| 1977 | Partnership returns, 1977, pages 29 & 47                            | Partnerships                           |
| 1978 | Sole Proprietorships returns, 1978, pages 27ff.                     | Sole proprietorships                   |
| 1978 | Partnership returns, 1978, pages 27ff.                              | Partnerships                           |
| 1979 | Sole Proprietorships returns, 1979-1780, pages 34ff.                | Sole proprietorships                   |
| 1979 | Partnership returns, 1979, pages 29ff.                              | Partnerships                           |
| 1980 | Sole Proprietorships returns, 1979-1780, pages 165ff.               | Sole proprietorships                   |
| 1980 | Partnership returns, 1980, pages 33ff.                              | Partnerships                           |
| 1998 | IRS Website, Integrated Business Data, Table 2                      | All; inconsistency in data for Finance |
| 1999 | IRS Website, Integrated Business Data, Table 2                      | All; inconsistency in data for Finance |
| 2000 | IRS Website, Integrated Business Data, Table 2                      | All; inconsistency in data for Finance |
| 2001 | IRS Website, Integrated Business Data, Table 2                      | All; inconsistency in data for Finance |
| 2002 | IRS Website, Integrated Business Data, Table 2                      | All; inconsistency in data for Finance |
| 2003 | IRS Website, Integrated Business Data, Table 2                      | All; inconsistency in data for Finance |

Notes: This table shows the sources for the tabulations of size bins for noncorporations.

## IA2.2 BEA Data

**Investment composition from BEA fixed asset tables** The BEA fixed asset tables report the investment composition by industry on an annual basis since 1901. There are 39 types of equipment, 31 types of structures, and 25 types of intellectual property. We include asset codes starting with “EP1” (computing equipment), “ENS” (software), and “RD” (R&D) in the numerator, and investment in all categories in the denominator. We match BEA sectors to our main sectors and subsectors, following Table [IA8](#). We drop 5210 Federal Reserve Banks in BEA fixed asset tables.

**Industry output from national accounts** We also use industry gross output from the BEA. Tables [IA9](#) and [IA10](#) show the mapping between industries in NIPA and our main sectors and subsectors. We do not reassign different components of “Information” and we do not reassign “Waste management and remediation services” to “Utilities: Electric and Gas” because detailed breakdown for these industries was not available from 1947 to 1962.



Table IA8 – Industry Mapping with BEA Fixed Asset Tables

| BEA Industry Name                                                  | BEA Code | Main Sector   | Subsector                     |
|--------------------------------------------------------------------|----------|---------------|-------------------------------|
| Farms                                                              | 110C     | Agriculture   |                               |
| Forestry, fishing, and related activities                          | 113F     | Agriculture   |                               |
| Oil and gas extraction                                             | 2110     | Mining        | Mining: Oil and Gas           |
| Mining, except oil and gas                                         | 2120     | Mining        | Mining: Other                 |
| Support activities for mining                                      | 2130     | Mining        | Mining: Oil and Gas           |
| Utilities                                                          | 2200     | Utilities     | Utilities: Electric and Gas   |
| Construction                                                       | 2300     | Construction  |                               |
| Wood products                                                      | 3210     | Manufacturing | Manufacturing: Wood           |
| Nonmetallic mineral products                                       | 3270     | Manufacturing | Manufacturing: Stone          |
| Primary metals                                                     | 3310     | Manufacturing | Manufacturing: Metals         |
| Fabricated metal products                                          | 3320     | Manufacturing | Manufacturing: Metals         |
| Machinery                                                          | 3330     | Manufacturing | Manufacturing: Machinery      |
| Computer and electronic products                                   | 3340     | Manufacturing | Manufacturing: Electrical     |
| Electrical equipment, appliances, and components                   | 3350     | Manufacturing | Manufacturing: Electrical     |
| Motor vehicles, bodies and trailers, and parts                     | 336M     | Manufacturing | Manufacturing: Transportation |
| Other transportation equipment                                     | 336O     | Manufacturing | Manufacturing: Transportation |
| Furniture and related products                                     | 3370     | Manufacturing | Manufacturing: Wood           |
| Miscellaneous manufacturing                                        | 338A     | Manufacturing | Manufacturing: Other          |
| Food, beverage, and tobacco products                               | 311A     | Manufacturing | Manufacturing: Food           |
| Textile mills and textile product mills                            | 313T     | Manufacturing | Manufacturing: Apparel        |
| Apparel and leather and allied products                            | 315A     | Manufacturing | Manufacturing: Apparel        |
| Paper products                                                     | 3220     | Manufacturing | Manufacturing: Paper          |
| Printing and related support activities                            | 3230     | Manufacturing | Manufacturing: Printing       |
| Petroleum and coal products                                        | 3240     | Manufacturing | Manufacturing: Chemicals      |
| Chemical products                                                  | 3250     | Manufacturing | Manufacturing: Chemicals      |
| Plastics and rubber products                                       | 3260     | Manufacturing | Manufacturing: Plastics       |
| Wholesale trade                                                    | 4200     | Trade         | Trade: Wholesale              |
| Retail trade                                                       | 44RT     | Trade         | Trade: Retail                 |
| Air transportation                                                 | 4810     | Utilities     | Utilities: Transportation     |
| Railroad transportation                                            | 4820     | Utilities     | Utilities: Transportation     |
| Water transportation                                               | 4830     | Utilities     | Utilities: Transportation     |
| Truck transportation                                               | 4840     | Utilities     | Utilities: Transportation     |
| Transit and ground passenger transportation                        | 4850     | Utilities     | Utilities: Transportation     |
| Pipeline transportation                                            | 4860     | Utilities     | Utilities: Transportation     |
| Other transportation and support activities                        | 487S     | Utilities     | Utilities: Transportation     |
| Warehousing and storage                                            | 4930     | Utilities     | Utilities: Transportation     |
| Publishing industries (including software)                         | 5110     | Manufacturing | Manufacturing: Printing       |
| Motion picture and sound recording industries                      | 5120     | Services      | Services: Entertainment       |
| Broadcasting and telecommunications                                | 5130     | Utilities     | Utilities: Communication      |
| Information and data processing services                           | 5140     | Services      | Services: Business            |
| Federal Reserve banks                                              | 5210     |               |                               |
| Credit intermediation and related activities                       | 5220     | Finance       | Finance: Banking              |
| Securities, commodity contracts, and investments                   | 5230     | Finance       | Finance: Banking              |
| Insurance carriers and related activities                          | 5240     | Finance       | Finance: Insurance            |
| Funds, trusts, and other financial vehicles                        | 5250     | Finance       | Finance: Holding Companies    |
| Real estate                                                        | 5310     | Finance       | Finance: Real Estate          |
| Rental and leasing services                                        | 5320     | Finance       | Services: Other               |
| Legal services                                                     | 5411     | Services      | Services: Business            |
| Computer systems design and related services                       | 5415     | Services      | Services: Business            |
| Miscellaneous professional, scientific, and technical services     | 5412     | Services      | Services: Business            |
| Management of companies and enterprises                            | 5500     | Finance       | Finance: Holding Companies    |
| Administrative and support services                                | 5610     | Services      | Services: Other               |
| Waste management and remediation services                          | 5620     | Services      | Services: Other               |
| Educational services                                               | 6100     | Services      | Services: Other               |
| Ambulatory health care services                                    | 6210     | Services      | Services: Other               |
| Hospitals                                                          | 622H     | Services      | Services: Other               |
| Nursing and residential care facilities                            | 6230     | Services      | Services: Other               |
| Social assistance                                                  | 6240     | Services      | Services: Other               |
| Performing arts, spectator sports, museums, and related activities | 711A     | Services      | Services: Entertainment       |
| Amusements, gambling, and recreation industries                    | 7130     | Services      | Services: Entertainment       |
| Accommodation                                                      | 7210     | Services      | Services: Hotels              |
| Food services and drinking places                                  | 7220     | Trade         | Trade: Retail: Restaurants    |
| Other services, except government                                  | 8100     | Services      | Services: Personal            |

Notes: This table shows the mapping between BEA industries and main sectors and subsectors in our data.

Table IA9 – Industry Mapping with NIPA: Pre-1997

| NIPA Industry Name                                                   | Main Sector   | Subsector                     |
|----------------------------------------------------------------------|---------------|-------------------------------|
| <b>Private industries</b>                                            |               |                               |
| <b>Agriculture, forestry, fishing, and hunting</b>                   | Agriculture   |                               |
| <b>Mining</b>                                                        | Mining        |                               |
| Oil and gas extraction                                               |               | Mining: Oil and Gas           |
| Mining, except oil and gas                                           |               | Mining: Other                 |
| Support activities for mining                                        |               | Mining: Oil and Gas           |
| <b>Utilities</b>                                                     | Utilities     | Utilities: Electric and Gas   |
| <b>Construction</b>                                                  | Construction  |                               |
| <b>Manufacturing</b>                                                 | Manufacturing |                               |
| Wood products                                                        |               | Manufacturing: Wood           |
| Nonmetallic mineral products                                         |               | Manufacturing: Stone          |
| Primary metals                                                       |               | Manufacturing: Metals         |
| Fabricated metal products                                            |               | Manufacturing: Metals         |
| Machinery                                                            |               | Manufacturing: Machinery      |
| Computer and electronic products                                     |               | Manufacturing: Electrical     |
| Electrical equipment, appliances, and components                     |               | Manufacturing: Electrical     |
| Motor vehicles, bodies and trailers, and parts                       |               | Manufacturing: Transportation |
| Other transportation equipment                                       |               | Manufacturing: Transportation |
| Furniture and related products                                       |               | Manufacturing: Wood           |
| Miscellaneous manufacturing                                          |               | Manufacturing: Other          |
| Food and beverage and tobacco products                               |               | Manufacturing: Food           |
| Textile mills and textile product mills                              |               | Manufacturing: Apparel        |
| Apparel and leather and allied products                              |               | Manufacturing: Apparel        |
| Paper products                                                       |               | Manufacturing: Paper          |
| Printing and related support activities                              |               | Manufacturing: Printing       |
| Petroleum and coal products                                          |               | Manufacturing: Chemicals      |
| Chemical products                                                    |               | Manufacturing: Chemicals      |
| Plastics and rubber products                                         |               | Manufacturing: Plastics       |
| <b>Wholesale trade</b>                                               | Trade         | Trade: Wholesale              |
| <b>Retail trade</b>                                                  | Trade         | Trade: Retail                 |
| <b>Transportation and warehousing</b>                                | Utilities     | Utilities: Transportation     |
| <b>Information</b>                                                   | Utilities     | Utilities: Communication      |
| <b>Finance and insurance</b>                                         | Finance       |                               |
| Federal Reserve banks, credit intermediation, and related activities |               | Finance: Banking              |
| Securities, commodity contracts, and investments                     |               | Finance: Banking              |
| Insurance carriers and related activities                            |               | Finance: Insurance            |
| Funds, trusts, and other financial vehicles                          |               | Finance: Holding Companies    |
| Real estate                                                          | Finance       | Finance: Real Estate          |
| Rental and leasing services and lessors of intangible assets         | Finance       | Services: Other               |
| <b>Professional, scientific, and technical services</b>              | Services      | Services: Business            |
| <b>Management of companies and enterprises</b>                       | Finance       | Finance: Holding Companies    |
| <b>Administrative and waste management services</b>                  | Services      | Services: Other               |
| <b>Educational services, health care, and social assistance</b>      | Services      | Services: Other               |
| <b>Arts, entertainment, and recreation</b>                           | Services      | Services: Entertainment       |
| Accommodation                                                        | Services      | Services: Hotels              |
| Food services and drinking places                                    | Trade         | Trade: Retail: Restaurants    |
| <b>Other services, except government</b>                             | Services      | Services: Personal            |

Notes: This table shows the mapping between industries in NIPA before 1997 (first column) and main sectors and subsectors in our data (second and third columns).

Table IA10 – Industry Mapping with NIPA: Post-1997

| NIPA Industry Name                                                   | Main Sector   | Subsector                     |
|----------------------------------------------------------------------|---------------|-------------------------------|
| <b>Private industries</b>                                            |               |                               |
| <b>Agriculture, forestry, fishing, and hunting</b>                   | Agriculture   |                               |
| <b>Mining</b>                                                        | Mining        |                               |
| Oil and gas extraction                                               |               | Mining: Oil and Gas           |
| Mining, except oil and gas                                           |               | Mining: Other                 |
| Support activities for mining                                        |               | Mining: Oil and Gas           |
| <b>Utilities</b>                                                     | Utilities     | Utilities: Electric and Gas   |
| <b>Construction</b>                                                  | Construction  |                               |
| <b>Manufacturing</b>                                                 | Manufacturing |                               |
| Wood products                                                        |               | Manufacturing: Wood           |
| Nonmetallic mineral products                                         |               | Manufacturing: Stone          |
| Primary metals                                                       |               | Manufacturing: Metals         |
| Fabricated metal products                                            |               | Manufacturing: Metals         |
| Machinery                                                            |               | Manufacturing: Machinery      |
| Computer and electronic products                                     |               | Manufacturing: Electrical     |
| Electrical equipment, appliances, and components                     |               | Manufacturing: Electrical     |
| Motor vehicles, bodies and trailers, and parts                       |               | Manufacturing: Transportation |
| Other transportation equipment                                       |               | Manufacturing: Transportation |
| Furniture and related products                                       |               | Manufacturing: Wood           |
| Miscellaneous manufacturing                                          |               | Manufacturing: Other          |
| Food and beverage and tobacco products                               |               | Manufacturing: Food           |
| Textile mills and textile product mills                              |               | Manufacturing: Apparel        |
| Apparel and leather and allied products                              |               | Manufacturing: Apparel        |
| Paper products                                                       |               | Manufacturing: Paper          |
| Printing and related support activities                              |               | Manufacturing: Printing       |
| Petroleum and coal products                                          |               | Manufacturing: Chemicals      |
| Chemical products                                                    |               | Manufacturing: Chemicals      |
| Plastics and rubber products                                         |               | Manufacturing: Plastics       |
| <b>Wholesale trade</b>                                               | Trade         | Trade: Wholesale              |
| <b>Retail trade</b>                                                  | Trade         | Trade: Retail                 |
| <b>Transportation and warehousing</b>                                | Utilities     | Utilities: Transportation     |
| <b>Information</b>                                                   | Utilities     | Utilities: Communication      |
| <b>Finance, insurance, real estate, rental, and leasing</b>          | Finance       |                               |
| Federal Reserve banks, credit intermediation, and related activities |               | Finance: Banking              |
| Securities, commodity contracts, and investments                     |               | Finance: Banking              |
| Insurance carriers and related activities                            |               | Finance: Insurance            |
| Funds, trusts, and other financial vehicles                          |               | Finance: Holding Companies    |
| <b>Real estate and rental and leasing</b>                            |               | Finance: Real Estate          |
| <b>Professional, scientific, and technical services</b>              | Services      | Services: Business            |
| <b>Management of companies and enterprises</b>                       | Finance       | Finance: Holding Companies    |
| <b>Administrative and waste management services</b>                  | Services      | Services: Other               |
| <b>Educational services, health care, and social assistance</b>      | Services      | Services: Other               |
| <b>Arts, entertainment, and recreation</b>                           | Services      | Services: Entertainment       |
| Accommodation                                                        | Services      | Services: Hotels              |
| Food services and drinking places                                    | Trade         | Trade: Retail: Restaurants    |
| <b>Other services, except government</b>                             | Services      | Services: Personal            |

Notes: This table shows the mapping between industries in NIPA after 1997 (first column) and main sectors and subsectors in our data (second and third columns).

## IA2.3 Variable Construction

In the following we explain our variable construction process.

### SOI Data

#### 1. Top $x\%$ share among U.S. corporations

##### Definition

- Top  $x\%$  asset share = Assets of top  $x\%$  corporations by assets in a given year/Total assets of U.S. corporations in a given year. Used in Figure 1, Panel A of Figure 3, Panel A of Figure 4, Figure 5, Figure IA1, Figure IA2, Figure IA3, Figure IA15, Figure IA16, Table 3, Table 4, Table 6, Table IA1, Table IA2, Table IA3, Table IA4, and Table IA5.
- Top  $x\%$  receipt share = Receipts of top  $x\%$  corporations by receipts in a given year/Total receipts of U.S. corporations in a given year. Used in Figure 1, Panel A of Figure 3, Panel B of Figure 4, and Figure 6.
- Top  $x\%$  net income share = Net income of top  $x\%$  corporations by net income among corporations with positive net income in a given year/Total net income of U.S. corporations with positive net income in a given year. Used in Figure 1 and Panel A of Figure 3.

We digitize SOI tables of U.S. corporations by size bins such as the examples shown in Table 1. We can obtain top business shares for the aggregate economy and for different levels of aggregation, using SOI corresponding tables.

**Procedure** We examine several methods to estimate top shares. Results for these methods are over 0.99 correlated as explained in Section 2 and shown in Figure IA1. The generalized Pareto interpolation is our default method.

- (a) We use the generalized Pareto interpolation method explained in Blanchet, Fournier, and Piketty (2022). The method first calculates the inverted Pareto coefficient  $b(p_i)$  for each threshold  $i$  where  $p_i$  is the fraction of businesses with assets (receipts/net income) more than  $y_i$ , and  $b(p_i)$  is the ratio between the average assets (receipts/net income) above  $y_i$  and the threshold  $y_i$ . It then derives a continuous curve of inverted Pareto coefficients, conditional on the information from the tabulation.
- (b) We fit lognormal curves to the size bins and interpolate the lognormal curves as explained in Internet Appendix IA2.4.
- (c) We can also directly add up the top bins such that the number of businesses in these bins approximates  $x\%$ , as long as the top bin contains less than  $x\%$  of businesses. Specifically, if

the total number of businesses is  $N$  and the number of businesses in the top  $k$  bins adds up to less than  $0.01N$  (whereas the top  $k + 1$  bins add up to more than  $0.01N$ ), then we take all the businesses in the top  $k$  bins and add  $(0.01N - \sum_{i=1}^k n_i)/n_{k+1}$  fraction from the  $k + 1$ th bin (where  $n_i$  denotes the number of businesses in the  $i$ th bin). In other words, we take all businesses in the top  $k$  bins and fill in the residual from the  $k + 1$ th bin.

## 2. Relative concentration among top businesses

### Definition

- Top  $x\%$  asset share among top  $y\%$  = Assets of top  $x\%$  corporations by assets in a given year/Assets of top  $y\%$  corporations by assets in a given year. Used in Figure 2, Panel B of Figure 3, and Figure 5.
- Top  $x\%$  receipt share among top  $y\%$  = Receipts of top  $x\%$  corporations by receipts in a given year/Receipts of top  $y\%$  corporations by receipts in a given year. Used in Figure 2.
- Top  $x\%$  net income share among top  $y\%$  = Net income of top  $x\%$  corporations by net income among corporations with positive net income in a given year/Net income of top  $y\%$  corporations by net income among corporations with positive net income in a given year. Used in Figure 2.

**Procedure** For each metric among assets, receipts, and net income, we obtain top  $x\%$  and top  $y\%$  shares as explained above, and take the ratio of top  $x\%$  share/top  $y\%$  share.

## 3. Top business receipt share among corporations plus noncorporations

**Definition** Top  $x\%$  receipt share among all businesses (direct estimate) = Receipts of top  $x\%$  businesses by receipts (among corporations and noncorporations) in a given year/Total receipts of corporations and noncorporations in a given year. Used in Figure 6.

**Procedure** We estimate this ratio in years where we have tabulations of noncorporations by size bins based on receipts. We first estimate separate distributions for corporations, partnerships, and sole proprietorships, and then merge these three distributions using the `gpinter` routine of Blanchet, Fournier, and Piketty (2022) to obtain estimates for top receipt shares among all businesses.

## 4. Top $N$ corporations' share among corporations

### Definition

- Top  $N$  asset share = Assets of top  $N$  corporations by assets in a given year/Total assets of U.S. corporations in a given year. Used in Figure 7 and Figure IA5.

- Top  $N$  receipt share = Receipts of top  $N$  corporations by receipts in a given year/Total receipts of U.S. corporations in a given year. Used in Figure 7, Figure IA5, and Figure IA6.
- Top  $N$  net income share = Net income of top  $N$  corporations by net income among corporations with positive net income in a given year/Total net income of U.S. corporations with positive net income in a given year. Used in Figure 7 and Figure IA5.

**Procedure** We use the generalized Pareto interpolation method described above.

#### 5. Top $N$ corporations' share among corporations plus noncorporations

##### **Definition**

- Top  $N$  asset share = Assets of top  $N$  corporations by assets in a given year/Estimated total assets of corporations plus noncorporations in a given year. Used in Figure 7 and Figure IA5.
- Top  $N$  receipt share = Receipts of top  $N$  corporations by receipts in a given year/Total receipts of corporations plus noncorporations in a given year. Used in Figure 7, Figure IA5, and Figure IA6.
- Top  $N$  net income share = Net income of top  $N$  corporations by net income among corporations with positive net income in a given year/Estimated total net income of corporations plus noncorporations (with positive net income) in a given year. Used in Figure 7 and Figure IA5.

**Procedure** For the top  $N$  asset share, we obtain data on total nonfinancial corporate and noncorporate assets from the Financial Accounts of the United States. We multiply the top  $N$  corporate asset share with the ratio of nonfinancial corporate assets to nonfinancial business assets (corporate plus noncorporate) from the Financial Accounts. We exclude farm assets and residential real estate to better match SOI data (this adjustment mainly affects the level of the top  $N$  asset share but not the time trend).

For the top  $N$  receipt share, we use data from SOI on noncorporations and multiply the top  $N$  corporate receipt share with the ratio of corporate to noncorporate receipts. For the top  $N$  net income share, we use data from SOI on noncorporations and multiply the top  $N$  corporate net income share with the ratio of corporate to noncorporate net income (among businesses with positive net income).

#### 6. Sales share of top 20 corporations by assets to compare with census concentration ratios

**Definition** Top 20 sales share in SOI data = Sales of top 20 corporations by assets/Total sales of corporations. Used in Panel B of Figure IA7 (for these granular industries the SOI tabulations only provide size bins by assets).

**Procedure** For each industry, we add up the top  $k$  bins and  $\alpha$  fraction of the  $k + 1$ th bin, such as the number of businesses in the top  $k$  bins plus  $\alpha$  fraction of the number of businesses in the  $k + 1$ th bin is 20. We sum the sales of businesses in the top  $k$  bins and  $\alpha$  fraction of the sales of the  $k + 1$ th bin for the numerator. (Here we cannot use the Pareto interpolation because we need the sales of the top businesses by *assets*, given that only size bins by assets are available for these granular industries.)

#### 7. Top 1% share including international subsidiaries of U.S. corporations

**Definition** Adjusted top 1% asset share = (Assets of top 1% corporations by assets + estimated assets of their international affiliates)/(Total assets of U.S. corporations + total assets of the foreign affiliates of U.S. multinationals). Used in Figure [IA9](#).

**Procedure** We can obtain data on the assets of the foreign affiliates of U.S. multinationals using Activities of U.S. Multinational Enterprises from the BEA. The data are available by the industry of the U.S. parent, though only for the assets (not sales) of international affiliates. For the denominator in the ratio above, we directly add the assets of the foreign affiliates to the total assets of U.S. corporations in the SOI data. For the numerator, we use two options: 1) assume that all international assets belong to the top businesses in the numerator (i.e., add all international assets to the numerator); 2) assume the top businesses' share of international assets is the same as their share of domestic assets (i.e., add a fraction of international assets to the numerator which is equal to the baseline top business shares among U.S. corporations).

#### 8. Sales concentration excluding exports

**Definition** Top  $x\%$  receipt share excluding exports = (Receipts of top  $x\%$  corporations by receipts – total exports)/(Total receipts of U.S. corporations – total exports). Used in Figure [IA17](#).

**Procedure** We obtain total exports from NIPA. To be conservative, we assume that all exports come from top businesses and subtract total exports from the numerator.

#### 9. Profitability (net income before tax/sales)

##### **Definition**

- Profitability among top  $x\%$  corporations by assets = Net income before tax among top  $x\%$  corporations by assets in a given year/Total receipts among top  $x\%$  corporations by assets in a given year. Used in Figure [8](#).
- Profitability among non-top  $x\%$  corporations by assets = Net income before tax among non-top  $x\%$  corporations by assets in a given year/Total receipts among non-top  $x\%$  corporations by assets in a given year. Used in Figure [8](#).

- Profitability among all corporations (SOI) = Net income before tax among all U.S. corporations in a given year/Total receipts among all U.S. corporations in a given year (SOI). Used in Figure IA11.
- Profitability among all corporations (BEA) = BEA corporate profits before tax among in a given year/Total receipts among all U.S. corporations in a given year (SOI). Used in Figure IA11.

**Procedure** For the last two ratios, we can directly obtain data on net income before tax and total receipts from SOI, as well as corporate profits before tax from BEA (where BEA uses economic depreciation instead of tax depreciation). For the first two ratios, we obtain net income before tax and receipts for top  $x\%$  businesses (and the rest) as follows. The SOI tables provide balance sheet and income statement information for each size bin (e.g., sales, net income, cash, inventory, fixed asset, equity, debt, among many other things), and Panel B of Table 1 shows a partial example. We add up the top bins such that the number of businesses in these bins approximates  $x\%$ , as long as the top bin contains less than  $x\%$  of businesses. For instance, if the total number of businesses is  $N$  and the number of businesses in the top  $k$  bins adds up to less than  $0.01N$  (whereas the top  $k + 1$  bins add up to more than  $0.01N$ ), then we take all the businesses in the top  $k$  bins and add  $(0.01N - \sum_{i=1}^k n_i)/n_{k+1}$  fraction from the  $k + 1$ th bin (where  $n_i$  denotes the number of businesses in the  $i$ th bin). We do so to estimate the net income and receipts of top  $x\%$  businesses, and the remaining amount of net income and receipts belong to non-top  $x\%$  businesses.

#### 10. Fixed assets/total assets

##### Definition

- Fixed assets over total assets among top  $x\%$  corporations by assets = Fixed assets among top  $x\%$  corporations by assets in a given year/Total assets among top  $x\%$  corporations by assets in a given year. Used in Figure IA12.
- Fixed assets over total assets among non-top  $x\%$  corporations by assets = Fixed assets among non-top  $x\%$  corporations by assets in a given year/Total assets among non-top  $x\%$  corporations by assets in a given year. Used in Figure IA12.

**Procedure** The SOI tables provide balance sheet and income statement information for each size bin, as can be seen from the examples in Table 1. We add up the top bins such that the number of businesses in these bins approximates  $x\%$ , as long as the top bin contains less than  $x\%$  of businesses. For instance, if the total number of businesses is  $N$  and the number of businesses in the top  $k$  bins adds up to less than  $0.01N$  (whereas the top  $k + 1$  bins add up to more than  $0.01N$ ), then we take all the businesses in the top  $k$  bins and add  $(0.01N - \sum_{i=1}^k n_i)/n_{k+1}$  fraction from the  $k + 1$ th bin (where  $n_i$  denotes the number of businesses in the  $i$ th bin). We do so to estimate



the fixed assets and total assets of top  $x\%$  businesses, and the remaining amount of fixed assets and total assets belong to non-top  $x\%$  businesses.

## 11. Book equity/total assets

### Definition

- Book equity over total assets among top  $x\%$  corporations by assets = Book equity among top  $x\%$  corporations by assets in a given year/Total assets among top  $x\%$  corporations by assets in a given year. Used in Figure IA19.
- Book equity over total assets among non-top  $x\%$  corporations by assets = Book equity among non-top  $x\%$  corporations by assets in a given year/Total assets among non-top  $x\%$  corporations by assets in a given year. Used in Figure IA19.

**Procedure** The SOI tables provide balance sheet and income statement information for each size bin, as can be seen from the examples in Table 1. We add up the top bins such that the number of businesses in these bins approximates  $x\%$ , as long as the top bin contains less than  $x\%$  of businesses. For instance, if the total number of businesses is  $N$  and the number of businesses in the top  $k$  bins adds up to less than  $0.01N$  (whereas the top  $k + 1$  bins add up to more than  $0.01N$ ), then we take all the businesses in the top  $k$  bins and add  $(0.01N - \sum_{i=1}^k n_i)/n_{k+1}$  fraction from the  $k + 1$ th bin (where  $n_i$  denotes the number of businesses in the  $i$ th bin). We do so to estimate the book equity and total assets of top  $x\%$  businesses, and the remaining amount of book equity and total assets belong to non-top  $x\%$  businesses.

## Other Data

### 1. Sales concentration in census data

#### Definition

- Average CR20 among manufacturing industries: equal weighted or sales-weighted average of CR20 among four-digit manufacturing SIC codes before 1997 and six-digit manufacturing NAICS codes after 1997. Used in Panel A of Figure IA7.
- CR20 among manufacturing and non-manufacturing industries in 2012. Used in Panel B of Figure IA7.

**Procedure** We use data from the census website for 1947 to 1992 (manufacturing at four-digit SIC level) and 1997 onward (manufacturing and non-manufacturing at two-digit to six-digit NAICS level):

<https://www.census.gov/data/tables/1992/econ/census/concentraion-ratio-data.html>, <https://data.census.gov/>.

## 2. Top business share using Compustat data

### Definition

- Compustat top 500 assets share = Total assets of top 500 Compustat firms by assets/Total corporate assets. Used in Figure [IA8](#).
- Compustat top 500 sales share = Total sales of top 500 Compustat firms by sales/Total corporate receipts. Used in Figure [IA8](#).

**Procedure** We sum the assets (sales) of top 500 Compustat firms by assets (sales), and divide by total corporate assets (receipts) from SOI.

## 3. Employment concentration in census BDS data

**Definition** Top  $x\%$  employment share = Employment of top  $x\%$  firms by employment in a given year/Total employment in a given year. Used in Figure [IA10](#).

**Procedure** We use annual census BDS tabulations of employment for size bins by employment and apply generalized Pareto interpolation to obtain estimates of employment concentration.

## 4. Investment rate

### Definition

- Investment rate of fixed assets = Investment in fixed assets/Stock of fixed assets. Used in Figure [IA13](#).
- Investment rate of fixed assets and intangibles = Investment in fixed assets and intangibles/Stock of fixed assets and intangibles. Used in Figure [IA13](#).

**Procedure** We use BEA fixed asset tables to obtain investment and stock of fixed assets and intangibles for each industry-year.

## 5. Entry rate

**Definition** Entry rate = New firms/Total firms. Used in Figure [IA14](#).

**Procedure** We use census BDS data to obtain new firms and total firms for each industry-year.

## 6. Investment in R&D and IT

### Definition

- Investment in R&D and IT in total investment = Investment in R&D and IT in BEA fixed asset tables/Total investment in BEA fixed asset tables. Used in Table [4](#) and Table [IA2](#).
- Investment in R&D and IT relative to receipts = Investment in R&D and IT in BEA fixed asset tables/Receipts in SOI. Used in Table [IA1](#).

**Procedure** For the numerator, we use the BEA fixed asset tables to obtain investment in R&D and IT in a given year and a given industry. R&D includes all R&D categories in the BEA fixed asset tables (labelled “RDxx”), except non-business categories “RD91” (Private universities and colleges) and “RD92” (Other nonprofit institutions). IT includes all software categories (labelled “ENSx”) and computer equipment categories (labelled “EP1A” to “EP1H”). For the denominator, the first ratio uses total investment in fixed assets and intellectual property in BEA fixed asset tables for a given year and a given industry; the second ratio uses receipts in SOI data for a given year and a given industry.

## 7. Fixed cost share

**Definition** Median annual estimated firm-level fixed cost share among Compustat firms in each subsector. Used in Table 5.

**Procedure** The firm’s variable cost share is estimated as:  $(\alpha_i \text{COGS} + \beta_i \text{SG\&A}) / (\text{COGS} + \text{SG\&A})$ , where  $\alpha_i$  ( $\beta_i$ ) is the coefficient of regressing log change of cost of goods sold (SG&A) on log change in sales for each industry  $i$  (the regression uses quarterly data from 1953 to 2013 because Compustat data only become available in the early 1950s). The firm’s fixed cost share is one minus the variable cost share. The weights  $\alpha_i$  and  $(\beta_i)$  capture the extent to which COGS and SG&A vary with sales, and correspondingly the extent to which they are variable. In this way we do not assume that COGS or SG&A represents entirely variable or fixed costs. We use each subsector as an industry  $i$  and take the median firm-level fixed cost share every year.

## IA2.4 Lognormal Interpolation

Below we explain our procedure for using generalized lognormal curves to estimate top shares from businesses by size bins. We first fit the lognormal curve for each bin, and then combine these lognormal curves to compute top shares.

**Fitting lognormal to each bin** For each of the discrete size bin thresholds  $0 = t_0 < t_1 < t_2 < \dots t_K$ , we calculate the cumulative density function (CDF) in the data up to that threshold,  $0 \leq F^*(t_k) \leq 1$ ). For each  $k$  between 1 and  $K$ , we fit a lognormal distribution for the bin  $[t_k, t_{k+1}]$  by targeting the CDFs up to the end-points of each bin. In other words, we fit  $(\mu_k, \sigma_k)$  such that:

$$\begin{aligned} F^*(t_k) &= \Phi\left(\frac{\log t_k - \mu_k}{\sigma_k}\right) \\ F^*(t_{k+1}) &= \Phi\left(\frac{\log t_{k+1} - \mu_k}{\sigma_k}\right), \end{aligned} \tag{IA1}$$

where  $\Phi$  is the standardized normal CDF. The above equation translates into two linear equations in  $\sigma_k$  and  $\mu_k$ , which can be solved exactly:

$$\begin{aligned}\Phi^{-1}(F^*(t_k))\sigma_k &= \log t_k - \mu_k \\ \Phi^{-1}(F^*(t_{k+1}))\sigma_k &= \log t_{k+1} - \mu_k.\end{aligned}\tag{IA2}$$

For the largest size bin (businesses with size above  $t_K$ ), we fit the lognormal parameters  $(\mu_K, \sigma_K)$  to target the CDF value up to  $t_K$  and the average size of business above  $t_K$ ,  $E^*[x \geq t_K]$ . In other words, for each proposal of  $\mu_K$ , we use the equation

$$\Phi^{-1}(F^*(t_K))\sigma_K = \log t_K - \mu_K\tag{IA3}$$

to pin down the implied  $\sigma_K$ , and then confirm whether:

$$\exp\left(\mu_K + \frac{1}{2}\sigma_K^2\right) \cdot \left(1 - \Phi\left(\frac{\log t_K - (\mu_K + \sigma_K^2)}{\sigma_K}\right)\right) = E^*[x \geq t_K].\tag{IA4}$$

Lastly, for the smallest bin (businesses with size less than  $t_1$ ), we use the same parameters as the interval for  $[t_1, t_2]$ .

**Computing top shares** From the above construction, we obtain that an interpolated lognormal CDF that is monotonically increasing (as long as  $F^*(t_k)$  in the raw data is monotonically increasing). Thus, to compute the top  $p$  share like  $p = 0.01$  (or other thresholds), we first need to compute  $t^*$  such that the interpolated CDF  $F_{Lognormal}$  is  $p$ :  $F_{Lognormal}(t^*) = p$ . This can be done by finding which size bin contains the cutoff for  $p$  (by comparing  $p$  to the CDF values in the data at each bin threshold), and then inverting the log-normal CDF at that interval to obtain the precise  $t^*$ .

Then, to compute the top share, it suffices to compute the (interpolated) mean conditional on firms with size  $x > t^*$ , and normalize that by the total mean, which we denote as  $\mu_0$ . For this, we use the fact that  $\int_a^b x dF(x)$  and  $\int_a^b dF(x)$  for  $F \sim \text{Log Normal}(\mu, \sigma)$  are respectively given by:

$$\begin{aligned}M(\mu, \sigma, a, b) &\equiv \exp\left(\mu + \frac{1}{2}\sigma^2\right) \cdot \left(\Phi\left(\frac{\log b - (\mu + \sigma^2)}{\sigma}\right) - \Phi\left(\frac{\log a - (\mu + \sigma^2)}{\sigma}\right)\right) \\ Z(\mu, \sigma, a, b) &\equiv \Phi\left(\frac{\log b - \mu}{\sigma}\right) - \Phi\left(\frac{\log a - \mu}{\sigma}\right).\end{aligned}\tag{IA5}$$

Suppose  $t_k < t^* < t_{k+1}$ . The conditional mean of the fitted distribution  $\hat{E}[x \geq t^*]$  is given by:

$$\hat{E}[x \geq t^*] = \frac{M(\mu_k, \sigma_k, t^*, t_{k+1}) + \sum_{j=k+1}^{K-1} M(\mu_j, \sigma_j, t_j, t_{j+1})}{Z(\mu_k, \sigma_k, t^*, t_{k+1}) + \sum_{j=k+1}^{K-1} Z(\mu_j, \sigma_j, t_j, t_{j+1})}.\tag{IA6}$$

If  $t^* > t_K$  (the largest bin has more than 1% businesses), the conditional mean of the fitted distribution

$\hat{E}[x \geq t^*]$  is given by:

$$\hat{E}[x \geq t^*] = \frac{M(\mu_K, \sigma_K, t^*, \infty)}{Z(\mu_K, \sigma_K, t^*, \infty)}. \quad (\text{IA7})$$

Our fitted top share is given by  $\frac{\hat{E}[x \geq t^*]}{\mu_0}$ , where  $\mu_0$  is the overall mean.

## IA3 Model

In the following, we present a simple model that illustrates the impact of scalable technologies as summarized in Section 4.1. We consider the existence of a traditional technology and the introduction of a new technology that decreases marginal costs but requires greater upfront spending. We show that technological development of this form will increase concentration and industry output. In addition, by allowing markups to be exogenous (e.g., [Covarrubias, Gutiérrez, and Philippon \(2020\)](#)), we clarify that the introduction of the new technology does not need to be accompanied by higher profitability.

### IA3.1 Static Case

#### IA3.1.1 Setup

We use the standard nested CES demand structure, where there is a continuum of firms in industry  $k$  indexed by  $i \in [0, N_k]$ . In other words, a firm  $i$  in industry  $k$  faces demand:

$$y_{i,k} = Y_k \cdot \left( \frac{p_{i,k}}{P_k} \right)^{-\sigma}, \quad (\text{IA8})$$

where  $p_{i,k}$  is the price, and  $P_k^{1-\sigma} = \int_0^{N_k} p_{i,k}^{1-\sigma} di$  is the aggregate price index for industry  $k$ , with  $N_k$  being the mass of firms in industry  $k$ . The aggregate demand for industry  $k$  is given by:

$$Y_k = \bar{Y} \left( \frac{P_k}{\bar{P}} \right)^{-\epsilon}, \quad (\text{IA9})$$

with the aggregate price index  $\bar{P}^{1-\epsilon} = \int_0^1 P_{k,t}^{1-\epsilon} dk$ . Appendix [IA3.1.4](#) shows the detailed CES aggregator that justifies the above demand function.

Firms pay an entry cost  $\kappa$  to enter the market (as in [Autor et al. \(2020\)](#), [Covarrubias, Gutiérrez, and Philippon \(2020\)](#), among others). After entry, each firm  $i$  observes its idiosyncratic productivity,  $a_i$ . Depending on the realization of its idiosyncratic productivity, a firm has three options:

1. **Exit immediately.**
2. **Operate with old technology:** Pay an upfront cost  $\phi$  and operate a constant-returns-to-scale technology with per-unit productivity  $a_i$  (or per-unit cost of  $1/a_i$ ). In other words, the firm uses  $L$  units of input (normalized to unit cost) to produce  $a_i L$  units of output.
3. **Operate with new technology:** Pay an upfront cost  $\Phi(h)$  and operate a constant-returns-to-scale technology with per-unit productivity  $A(a_i, h)$  (or per-unit cost of  $1/A(a_i, h)$ ). In other words, the firm uses  $L$  units of input (normalized to unit cost) to produce  $A(a_i, h)L$  units of output.

This choice between the new (old) technology with higher (lower) upfront costs and lower (higher) marginal costs is similar to the spirit of the model in [Hsieh and Rossi-Hansberg \(2022\)](#). For simplicity of illustration, firms that decide to stay will operate in perpetuity under the same per-period productivity

( $a_i$  for the old technology and  $A(a_i, h)$  for the new technology), with profits in each period discounted at a constant rate  $R$ .<sup>31</sup>

The parameter  $h \geq 1$  is an index of the scalability of the new technology. We examine how the development of the new technology affects concentration by comparing the equilibrium outcomes under  $h = 1$  (where the two technologies are one and the same) with those under  $h > 1$ . We assume that a)  $\Phi(h) \geq \phi$ : the new technology requires a greater upfront cost than the old technology, and b)  $A(a_i, h) \geq a_i$ : the new technology enables firms to produce each unit more efficiently. Furthermore, we assume  $\Phi'(h) > 0$ ,  $\frac{\partial A}{\partial h} > 0$ . For tractability, we assume a simple functional form for  $\Phi$  and  $A$ :  $\Phi(h) = h^\eta \phi$  and  $A(a_i, h) = h \cdot a_i$ , with  $\eta > 1$ .

Denote the time-0 profit of a firm with idiosyncratic productivity  $a_i$  using the old and new technology as  $\pi_t(a_i)$  and  $\pi'_t(a_i)$  respectively. Then, the net present value of each technology is given by:

$$\begin{aligned}\Pi(a_i) &= \underbrace{\sum_{t=1}^{\infty} \frac{1}{R^t} \pi_t(a_i)}_{\text{Discounted Profit}} - \underbrace{\phi}_{\text{Investment}}, \\ \Pi'(a_i) &= \underbrace{\sum_{t=1}^{\infty} \frac{1}{R^t} \pi'_t(a_i)}_{\text{Discounted Profit}} - \underbrace{\Phi(h)}_{\text{Investment}}.\end{aligned}\tag{IA10}$$

To solve for equilibrium entry, profits, and concentration, we make some simplifying assumptions. First, we assume exogenous markups (as in [Covarrubias, Gutiérrez, and Philippon \(2020\)](#)).

**Assumption IA1 (Exogenous markups).** *Firms adopt an exogenous markup  $\mu$ : a firm with constant returns to scale technology  $a_i$  has unit cost  $\frac{1}{a_i}$  and set  $p_i = \frac{1+\mu}{a_i}$ .*

We make this assumption to demonstrate that trends in concentration do not have to be accompanied by trends in profitability. The use of an exogenous markup allows us to flexibly allow any movements in markups. Even if we allow markups to be endogenously set at the profit-maximizing level  $\mu^* = \frac{1}{\sigma-1}$ , all of our conclusions remain.

Second, we assume free entry to pin down the number of firms  $N_k$ .

**Assumption IA2 (Free entry).** *The entry cost is equal to the ex ante expected net present value of the firm:*

$$\kappa = E_{a_i \sim F} [\max \{0, \Pi(a_i), \Pi'(a_i)\}].\tag{IA11}$$

Finally, we assume that the new technology requires a sufficiently high upfront cost, such that it does not completely dominate the pre-existing technology for all firms. Under our functional form assumption, the above assumption translates to the following condition:

**Assumption IA3 (Non-domination of technology).** *Let  $\Phi(h) = h^\eta \phi$  be the investment cost function. We assume  $\eta > \sigma - 1$ .*

---

<sup>31</sup>This shortcut allows us to illustrate the role of technological innovation on rising concentration without fully specifying a dynamic model.

### IA3.1.2 Solution

Under the above assumptions, the exogenous markup assumption pins down the demand for firm with the old technology: the firm in industry  $k$  with unit cost  $\frac{1}{a_i}$  charges a price of  $\frac{1+\mu}{a_i}$ , and thus faces demand:

$$y_{i,k} = Y_k \cdot \left( \frac{1+\mu}{a_i} \frac{1}{P_k} \right)^{-\sigma}. \quad (\text{IA12})$$

This pins down the input choice  $L_{i,k}^*$  at:

$$L_{i,k}^* = \frac{1}{a_i} y_{i,k} = \frac{1}{a_i} Y_k \cdot \left( \frac{1+\mu}{a_i} \frac{1}{P_k} \right)^{-\sigma}. \quad (\text{IA13})$$

The expressions for the firm that adopts the new technology follows similarly. Then, one can derive the following expression for  $\Pi$  and  $\Pi'$ :

$$\begin{aligned} \Pi(a_i) &= \frac{R}{R-1} \cdot \frac{\mu}{(1+\mu)^\sigma} Y_k \cdot P_k^\sigma a_i^{\sigma-1} - \phi, \\ \Pi'(a_i) &= \frac{R}{R-1} \cdot \frac{\mu}{(1+\mu)^\sigma} Y_k \cdot P_k^\sigma (h \cdot a_i)^{\sigma-1} - \phi \cdot h^\eta. \end{aligned} \quad (\text{IA14})$$

Given the above assumptions, one can show that there will be three groups of firms in equilibrium: 1) the most productive firms adopt the new technology, 2) the next productive firms operate with the old technology, and 3) the least productive firms exit immediately.

**Proposition IA1.** *In equilibrium, there exists two thresholds  $a^*$  and  $a^{**}$ , defined by:*

$$\begin{aligned} \Pi(a^*) = 0 &\iff \phi = \frac{R}{R-1} \frac{\mu}{(1+\mu)^\sigma} Y_k \cdot P_k^\sigma (a^*)^{\sigma-1}, \\ \Pi(a^{**}) = \Pi'(a^{**}) &\iff a^{**} = \left( \frac{h^\eta - 1}{h^{\sigma-1} - 1} \right)^{1/(\sigma-1)} a^*. \end{aligned} \quad (\text{IA15})$$

*In equilibrium, firms with  $a_i < a^*$  exit, firms with  $a^* \leq a_i \leq a^{**}$  use the old technology, and firms with  $a_i \geq a^{**}$  use the new technology. The thresholds  $a^*$  and  $a^{**}$  depend positively on the markup  $\mu$  and negatively on the discount rate  $R$ .*

Second, let  $S_t(a_i)$  be the per-period revenue, and  $\tilde{\pi}_t(a_i) = \frac{\max\{\pi_t(a_i), \pi'_t(a_i)\}}{S_t(a_i)}$  be the profitability of the firm with idiosyncratic productivity  $a_i$  in equilibrium. We can derive the following expressions for firms that choose to operate.

**Proposition IA2.** *Let  $dF^*$  be the (normalized) distribution of  $a_i$  conditional on  $a_i \geq a^*$ , and let  $A^*$  be given by:<sup>32</sup>*

$$A^* = \left( \int_{a^*}^{a^{**}} a^{\sigma-1} dF^*(a) + h^{\sigma-1} \int_{a^{**}}^{\infty} a^{\sigma-1} dF^*(a) \right)^{\frac{1}{\sigma-1}}. \quad (\text{IA16})$$

<sup>32</sup>In other words,  $A^*$  is the  $\sigma - 1$  norm of the productivity of firms in operation; it can be loosely interpreted as the “average” productivity.



Then,

$$S_t(a_i) = \begin{cases} \left(\frac{a_i}{A^*}\right)^{\sigma-1} \frac{P_k \cdot Y_k}{N_k} & a^* \leq a_i \leq a^{**}, \\ \left(\frac{h \cdot a_i}{A^*}\right)^{\sigma-1} \frac{(P_k \cdot Y_k)}{N_k} & a^{**} \leq a_i. \end{cases} \quad (\text{IA17})$$

Furthermore,  $\tilde{\pi}_t(a_i) = \frac{\mu}{1+\mu}$ : the profitability of a firm corresponds one-to-one with the exogenous markup  $\mu$ . In particular, it does not depend on the technology index  $h$ .

Proposition IA2 implies that profitability can be distinct from how technology affects concentration.

Finally, using the Pareto distribution assumption, we can derive the expression for industry concentration by sales. To align with our empirical results, we calculate the share of the top 1% firms in total sales. For simplicity, we assume that the parameters are such that the top 1% firms all belong to the group of firms that operate with the new technology.<sup>33</sup>

Then, the concentration measure is given by:

$$\zeta_{1\%} = \frac{h^{\sigma-1} \int_{\alpha a^*}^{\infty} a_i^{\sigma-1} dF^*(a)}{\int_{a^{**}}^{\infty} a_i^{\sigma-1} dF(a) + h^{\sigma-1} \int_{a^{**}}^{\infty} a_i^{\sigma-1} dF^*(a)}, \quad (\text{IA18})$$

where  $\alpha$  is a constant that only depends on the Pareto parameter  $k$ . Note that for aggregate sales to be finite, we need  $k > \sigma - 1$ . We can then obtain the concentration ratio.

**Proposition IA3.** *The concentration ratio (top 1% sales share) is given by:*

$$\zeta_{1\%} = C \cdot \frac{h^{\sigma-1}}{1 + (h^{\sigma-1} - 1)^{\frac{k}{\sigma-1}} (h^{\eta} - 1)^{1 - \frac{k}{\sigma-1}}}, \quad (\text{IA19})$$

where  $C$  is a constant independent of  $h$ .

### IA3.1.3 Comparative Statics

We consider a marginal increase in  $h$  from 1 (where the two technologies coincide) to  $h > 1$ . By taking the comparative statics of Equation (IA19) and Proposition IA2, we obtain the following result.

**Proposition IA4.** *A rise in  $h$  leads to greater industry concentration, as measured by the top 1% sales share. The share of upfront costs in total costs also increases with  $h$  for firms that adopt the new technology. Meanwhile, there is no change in the per-period profitability  $\tilde{\pi}$ , which depends on the exogenous markup  $\mu$ .*

Next, we shall examine the predictions for industry output given a marginal increase in  $h = 1 + \nu$  in one industry  $k$ . Due to the continuous CES setup, each industry is marginal and has no impact on the aggregate output, so the growth in industry output is the same as the growth in industry share.

**Proposition IA5.** *Assume  $\sigma > \epsilon > 1$  (the cross-industry elasticity is weaker than the within-industry elasticity) and  $\eta > \sigma - 1$  is sufficiently small.<sup>34</sup> Then, a rise in  $h$  leads to a rise in the industry's output and its share in the economy.*

<sup>33</sup>This holds as long as  $1 - F^*(a^{**}) > 0.01$ .

<sup>34</sup>Alternatively, one can assume that the technological improvement is sufficiently marginal, i.e.  $\nu \mapsto 0$ .

The first assumption is standard: it is easier for a consumer to substitute within a given industry than to substitute across industries. The second assumption requires that the rise in investment associated with the new technology is not prohibitively expensive. This reflects two opposing consequences of technological improvement on output: first, it increases the output for firms that use the new technology. This is the primary intuition behind the link between higher concentration and higher industry output. On the other hand, technological improvement crowds out the output of firms that do not use the new technology. This effect is typically second-order relative to the first effect, provided that the new technology does not require a prohibitive amount of investment.

In summary, our model provides a simple illustration in which technological development results in higher concentration as firms that use the new technology increase their output relative to other firms. Industry output also increases. Meanwhile, profitability does not have to change.

#### IA3.1.4 Details and Proofs

**CES setup** Recall the standard nested CES setup: let  $k$  be the index for the industry (ranging from 0 to 1), and let  $i \in [0, N_k]$  be the index for a firm in industry  $k$ . The standard nested CES model assumes that the goods are aggregated using the following aggregator:

$$Y_k^{\frac{\sigma-1}{\sigma}} = \int_0^{N_k} y_{i,k}^{\frac{\sigma-1}{\sigma}} di. \quad (\text{IA20})$$

The industry goods are also aggregated into a final consumption bundle:

$$\bar{Y} = \int_0^1 Y_k^{\frac{\epsilon-1}{\epsilon}} dk. \quad (\text{IA21})$$

The demand system then implies that there exists an industry price index:

$$P_k^{1-\sigma} = \int_0^{N_k} Y_{k,t} \left( \frac{p_{i,k}}{P_k} \right)^{-\sigma}, \quad (\text{IA22})$$

and an aggregate price index:

$$\bar{P}^{1-\epsilon} = \int_0^1 P_{k,t}^{1-\epsilon} dk. \quad (\text{IA23})$$

Given these price indices, industry and firm demands are given by:

$$\begin{aligned} Y_k &= \bar{Y} \left( \frac{P_k}{\bar{P}} \right)^{-\epsilon}, \\ y_{i,k} &= Y_{k,t} \left( \frac{p_{i,k}}{P_k} \right)^{-\sigma}. \end{aligned} \quad (\text{IA24})$$

**Proofs** For notational simplicity, we solve for the case as  $R \mapsto \infty$ : the general case of  $R$  follows identically.

*Proof of Proposition IA4.* To show that the concentration ratio, given by Equation (IA19), is increasing in  $h$ , we take the following approach. Recall that  $\frac{a^{**}}{a^*} = \left( \frac{h^\eta - 1}{h^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}}$  is both larger than 1 and increasing

in  $h$ . Denoting  $H = h^{\sigma-1} - 1$ , the above implies that

$$D(H) = \left( \frac{a^{**}}{a^*} \right)^{(\sigma-1)-k}$$

is smaller than 1 and is a decreasing function in  $H$ . Setting  $F(H) = \frac{H+1}{1+H \cdot D(H)}$ , it suffices to show that  $F$  is an increasing function of  $H$ . Differentiation yields:

$$\frac{\partial F}{\partial H} > 0 \iff 1 + H \cdot D(H) - (H+1) \cdot \left( D(H) + H \frac{\partial D(H)}{\partial H} \right) = 1 - D(H) - H D'(H) > 0,$$

which holds as we have already shown that  $D(H) < 1$  and  $D'(H) < 0$ .

Finally, for firms that adopt the new technology, their sales increase by a factor of  $h^{\sigma-1}$ . Given constant markup assumption, their total variable costs increase by the same factor. Meanwhile, their upfront costs increase by a factor of  $h^\eta$ , which is greater than  $h^{\sigma-1}$  by Assumption IA3. Thus, the share of upfront costs in total costs is an increasing function of  $h$  for firms that adopt the new technology.

Profitability does not rely on  $h$  given the fixed and exogenous markup assumption. The firm has unit cost  $\frac{1}{a_i}$  and charges a price of  $\frac{1+\mu}{a_i}$ , with profit of  $\frac{\mu}{a_i}$ , so profitability is  $\tilde{\pi} = \frac{\mu}{1+\mu}$ .  $\square$

*Proof of Proposition IA5.* The three equations are given as follows:

$$\begin{aligned} \phi &= \frac{\mu}{1+\mu} \cdot \left( \frac{a^*}{A^*} \right)^{\sigma-1} \frac{P_k Y_k}{N_k}, \\ P_k &= \frac{1+\mu}{N^{\frac{1}{\sigma-1}}} \frac{1}{A^*}, \\ \frac{\kappa}{\phi} &= (1 - F(a^*)) \left( \left( \frac{A^*}{a^*} \right)^{\sigma-1} - 1 \right). \end{aligned} \tag{IA25}$$

The first equation is the zero-profit condition for the least profitable firm that stays in operation (which determines  $a^*$ ). The second is by definition the aggregate price index for industry  $j$ . The final is the free-entry condition. Note that:

$$P_j Y_j \propto P_j^{1-\epsilon}. \tag{IA26}$$

Thus, for industry output to grow in  $h$ , it thus suffices to show that  $P_j(h)$  is decreasing in  $h$ . Combining the first two equations gives:

$$C = \frac{\mu}{(1+\mu)^\sigma} (a^*)^{\sigma-1} P_j^{\sigma-\epsilon}, \tag{IA27}$$

where  $C$  is a constant. Given our assumption  $\sigma > \epsilon$ , it thus suffices to show that  $a^*$  is rising with  $h$ . The third equation then implies that holding everything constant, we have that  $a^*$  rises if and only if  $\frac{A^*}{a^*}$  rises with  $h$ .

We can compute:

$$\begin{aligned} (A^*)^{\sigma-1} &= \int_{a^*}^{\infty} a^{\sigma-1} dF^*(a) + (h^{\sigma-1} - 1) \int_{a^{**}}^{\infty} a^{\sigma-1} dF^*(a) \\ &= E_*[a^{\sigma-1}] \cdot \left( 1 + (h^{\sigma-1} - 1) \left( \frac{a^{**}}{a^*} \right)^{(\sigma-1)-k} \right), \end{aligned} \tag{IA28}$$

where  $E_*$  is the expectation relative to  $F^*$ , which is a truncated Pareto distribution. By the properties of the distribution,  $E_*[a^{\sigma-1}] = (a^*)^{\sigma-1} \cdot D$ , where  $D$  is a constant (a function of the power-law parameter).

Therefore,  $A^*/a^*$  is increasing in  $h$  if and only if  $(h^{\sigma-1} - 1) \left(\frac{a^{**}}{a^*}\right)^{(\sigma-1)-k}$  is increasing in  $h$ . As the expression makes clear, there are two countervailing forces: an increase in  $h$  makes firms more productive ( $h^{\sigma-1} - 1$ ), but fewer firms adopt the technology, as the greater returns to scale technology only selects for the most efficient firms. We plug in:

$$a^{**}/a^* = \left( \frac{h^\eta - 1}{h^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}}, \quad (\text{IA29})$$

to get that  $(h^{\sigma-1} - 1) \left(\frac{a^{**}}{a^*}\right)^{(\sigma-1)-k}$  evaluates to:

$$(h^\eta - 1)^{1-\frac{k}{\sigma-1}} (h^{\sigma-1} - 1)^{\frac{k}{\sigma-1}}, \quad (\text{IA30})$$

with the log-derivative given by:

$$\left(1 - \frac{k}{\sigma-1}\right) \frac{\eta \cdot h^{\eta-1}}{h^\eta - 1} + \frac{k}{\sigma-1} \frac{(\sigma-1)h^{\sigma-2}}{h^{\sigma-1} - 1}. \quad (\text{IA31})$$

Recall that we have  $k > \sigma - 1$  (the moments need to be well-defined) and  $\eta > \sigma - 1$  (for there to be differential adoption of the technology). As  $h \mapsto 1$ , the second term dominates (and goes to  $\infty$ ), which leads to the expression in Equation (IA31) being positive, as desired. Alternatively, note that  $\frac{\eta \cdot h^{\eta-1}}{h^\eta - 1}$  is increasing in  $\eta$  (and consequently the first term is decreasing in  $\eta$ , and for  $\eta \mapsto \sigma - 1$ , Equation (IA31) converges to:

$$\frac{(\sigma-1)h^{\sigma-1}}{h^{\sigma-1} - 1} > 0, \quad (\text{IA32})$$

as desired. □

## IA3.2 Dynamic Model

We now present a dynamic extension of the model.

### IA3.2.1 Setup

We use the CES framework as before, with the per-period demand given by:

$$y_{i,t} = Y \cdot \left( \frac{p_{i,t}}{P_t} \right)^{-\sigma}, \quad (\text{IA33})$$

where  $P^{1-\sigma} = \int_0^N p_i^{1-\sigma} di$  and  $N$  is the mass of operating firms. As before, we assume that firms employ an exogenous markup  $\mu$ .

Now, however, the supply side is modified to allow for dynamics. In each time period  $t$ , a new technology ( $A_t$ ) arises, which enables firms with idiosyncratic productivity  $a_i$  to produce at constant returns to scale at efficiency  $A_t(a_i)$  (where  $A_t$  is an increasing function).<sup>35</sup> A firm has to invest  $\phi_t$  once to acquire the technology. The investment needs to be maintained (from the first period onwards) by paying a per-period cost  $f_t$ , which can be interpreted as a type of fixed operating costs (e.g., capital depreciation, overhead).

New firms have to pay an entry cost  $\kappa_t$  before choosing to enter. Once they enter, they discover their idiosyncratic productivity  $a_i$ , and choose whether to exit or operate by investing  $\phi_t$ .<sup>36</sup> We assume these firms have idiosyncratic productivity  $a_i$  drawn from  $F_t$ , which we assume to be fixed to the firm. The production function is then linear, i.e. constant-returns-to-scale, as is the case for the basic version of the model: firms with idiosyncratic productivity  $a_i$  operating under a technology  $A_s$  at time  $t$  produce  $A_t(a_i) \cdot L$  units of the output using  $L$  units of the input.

The per-period profit of the firm then is given by:

$$\pi_{i,t}(a, s) = \frac{\mu}{(1 + \mu)^\sigma} Y \cdot P_t^\sigma A_s(a_i)^{\sigma-1} - f_s. \quad (\text{IA34})$$

The price index  $P^{1-\sigma}$  now depends on the composition of the technology of the firms in operation. Let  $N_{t,s}$  be the mass of firms operating at time  $t$  that use the technology introduced at time  $s$ , with  $N_t = \sum_{s \leq t} N_{t,s}$ . Furthermore, denote the distribution of idiosyncratic productivity of those firms as  $F_{t,s}$ . Then, the aggregate price index is given by:

$$P_t^{1-\sigma} = \sum_{s \leq t} N_{t,s} \cdot \int \left( \frac{1 + \mu}{A_s(a)} \right)^{1-\sigma} dF_{t,s}(a). \quad (\text{IA35})$$

If we continue to set  $\bar{Y} = Y P_t$  (industry aggregate remains constant), then we have the following simplification. Combining the previous two equations, we obtain the following expressions for per period profit  $\pi_t$  and sales  $S_t$ :

$$\begin{aligned} \pi_t(a, s) &= \frac{\mu}{1 + \mu} \frac{\bar{Y}}{N_t} \left( \frac{A_s(a)}{A_t^*} \right)^{\sigma-1} - f_s, \\ S_t(a, s) &= \left( \frac{A_s(a)}{A_t^*} \right)^{\sigma-1} \frac{\bar{Y}}{N_t}, \end{aligned} \quad (\text{IA36})$$

where

$$A_t^* = \left( \sum_{s \leq t} \frac{N_{t,s}}{N_t} \int A_s^{\sigma-1}(a) dF_{t,s}(a) \right)^{\frac{1}{\sigma-1}}. \quad (\text{IA37})$$

<sup>35</sup>This notation assumes that the idiosyncratic productivity for each period satisfies a monotonicity property: if incumbent firm  $X$  is more productive than incumbent firm  $Y$  under the old technology, then it remains the case if both adopt the new technology. This simplification rules out technological leap-frogging.

<sup>36</sup>Here we rule out the adoption of an older technology for simplicity.

### IA3.2.2 Solving the Model

We make the following simplifying assumptions. First, we assume  $f_t = f$  (constant). Second, we observe that there is a one-to-one correspondence between the mass of future entrants  $N_{t,t}$  and the future trajectory of entry cost  $\kappa_t$ . Consequently, we can alternatively specify a sequence of the mass of future entrants  $N_{t,t}$ , which would imply a series of future entry costs.

Finally, to simplify the dynamic optimization problem each firm faces, we assume that each firms are myopic optimizers: they only seek to optimize the current period profits. We think of a period as roughly one decade (so this assumption is not unrealistic). We note that the dynamic problem faced by a firm with idiosyncratic productivity  $a_i$  using technology at time  $s$  is the same *regardless* of when the firm has entered the market. We parametrize the sequence of new technologies in the following way:  $A_t(a_i) = h^t \cdot a_i$ , and  $\phi_t = h^{\eta t} \cdot \phi$ , with  $F_t$  being the Pareto distribution with tail index  $k$ . In line with the one-to-one equivalence between  $N_{t,t}$  and  $\kappa_t$ , we specify  $N_{t,t} = \bar{N}$ .

The following describes the algorithm that specifies the general equilibrium. For each generation  $g = 1, 2, \dots, t$ , there are firms that use technology vintage  $g \leq s \leq t$ . The set of firms belonging to generation  $g$  that use technology  $s$  is given by those with idiosyncratic productivity belonging to a collection of intervals  $T_{g,s} = \cup_i (\ell_{g,s}^i, u_{g,s}^i)$ . The equilibrium is fully specified by  $T_{g,s}$ , which we continue to update with the introduction of a new technology.

Let  $\Psi_t = (A_t^*)^{\sigma-1} \cdot N_t$ . For companies using technology of vintage  $s$ , their per-period profits are given by:

$$\frac{\mu}{1 + \mu} \frac{\bar{Y}}{\Psi_t} \cdot (h^s a)^{\sigma-1} - f. \quad (\text{IA38})$$

On the other hand, the current period profit of adopting the new technology is given by:

$$\frac{\mu}{1 + \mu} \frac{\bar{Y}}{\Psi_t} \cdot (h^t a)^{\sigma-1} - f - \phi \cdot h^{\eta t}. \quad (\text{IA39})$$

Thus, for each firm using technology vintage  $s < t$ , the exit threshold  $\beta_{t,s}$  is given by:

$$\begin{aligned} \max \left\{ \frac{\mu}{1 + \mu} \frac{\bar{Y}}{\Psi_t} \cdot (h^t a)^{\sigma-1} - f - \phi \cdot h^{\eta t}, \frac{\mu}{1 + \mu} \frac{\bar{Y}}{\Psi_t} \cdot (h^s a)^{\sigma-1} - f \right\} &< 0 \\ \iff a < \left( \frac{1 + \mu}{\mu} \frac{\Psi_t}{\bar{Y}} \right)^{\frac{1}{\sigma-1}} \min \left\{ (f + \phi \cdot h^{\eta t})^{\frac{1}{\sigma-1}} h^{-t}, f^{\frac{1}{\sigma-1}} \cdot h^{-s} \right\} &= \beta_{t,s}. \end{aligned} \quad (\text{IA40})$$

Furthermore, the adoption threshold  $\gamma_{t,s}$  is given by:

$$\begin{aligned} \frac{\mu}{1 + \mu} \frac{\bar{Y}}{\Psi_t} \cdot (h^t a)^{\sigma-1} - f - \phi \cdot h^{\eta t} &> \frac{\mu}{1 + \mu} \frac{\bar{Y}}{\Psi_t} \cdot (h^s a)^{\sigma-1} - f \\ \iff a > \left( \frac{1 + \mu}{\mu} \frac{\Psi_t}{\bar{Y}} \right)^{\frac{1}{\sigma-1}} \cdot \left( \phi \cdot \frac{h^{\eta t}}{h^{t(\sigma-1)} - h^{s(\sigma-1)}} \right)^{\frac{1}{\sigma-1}} &= \gamma_{t,s}. \end{aligned} \quad (\text{IA41})$$

Finally, for each generation  $g$ , we record the minimum productivity of that generation at time  $g$ . In

other words, at time  $g$  when the generation  $g$  firms enter,  $T_{g,g} = (\alpha_g, \infty)$ .<sup>37</sup> For a given value of  $\Psi_t$ , we have:

$$\alpha_t = \left( \frac{1 + \mu}{\mu} \frac{\Psi_t}{\bar{Y}} (f + \phi \cdot h^{\eta t}) \right)^{\frac{1}{\sigma-1}} \cdot h^{-t}. \quad (\text{IA42})$$

Finally, the collection  $T_{g,s}$  for all  $g$  and  $s$  imply the true value of  $\Psi_t$ , given by the following formula:

$$\Psi_t = (A_t^*)^{\sigma-1} N_t = \sum_{g \leq t} \frac{k}{k+1-\sigma} \alpha_g^k N_{g,g} \left( \sum_{g \leq s \leq t} h^{s(\sigma-1)} \sum_i \left( (\ell_{g,s}^i)^{-(k+1-\sigma)} - (u_{g,s}^i)^{-(k+1-\sigma)} \right) \right) \quad (\text{IA43})$$

**Definition IA1.** A (myopic) dynamic equilibrium is given by the collection of  $\{\Psi_t, \alpha_t, N_{t,t}, T_{g,s}^t\}$  for each  $t \geq 1$ , such that  $T_{g,s}^t$  satisfy:

$$\begin{aligned} T_{g,s}^t &= (T_{g,s}^{t-1} \cap (\beta_{t,s}, \infty)) \cap (0, \gamma_{t,s}) \text{ for } g, s < t \\ T_{g,t}^t &= (T_{g,s}^{t-1} \cap (\beta_{t,s}, \infty)) \cap (\gamma_{t,s}, \infty) \text{ for } g < t \\ T_{t,t}^t &= (\alpha_t, \infty), \end{aligned} \quad (\text{IA44})$$

where  $\beta_{t,s}$ ,  $\gamma_{t,s}$ , and  $\alpha_{t,s}$  are given by Equations (IA40), (IA41), and (IA42), with  $\Psi_t$  conditional on  $T_{g,s}^t$  given by Equation (IA43).

Thus, we compute the dynamic equilibrium given the following algorithm:

1. Initialize the equilibrium for  $t = 1$ . Here, note that  $\Psi_1$  takes a relatively simple form:

$$\begin{aligned} \Psi_1 &= \bar{N} \frac{k}{k+1-\sigma} h^{\sigma-1} \alpha_1^{\sigma-1} \\ \alpha_1^{\sigma-1} &= \frac{1 + \mu}{\mu} \frac{\Psi_1}{\bar{Y}} (f + \phi h^{\eta}) h^{-(\sigma-1)}, \end{aligned} \quad (\text{IA45})$$

which implies:

$$\bar{N}_1 = \frac{k+1-\sigma}{k} \frac{\mu}{1+\mu} \frac{\bar{Y}}{f + \phi h^{\eta}}, \quad (\text{IA46})$$

and  $\alpha_1$  can be set to a constant 1.

2. Subsequently, for  $t > 1$ : have a record of  $T_{g,s}$  for  $g, s < t$ .
  - (a) Posit a value for  $\Psi_t$ .
  - (b) Update the implied  $T_{g,s}$  for  $g, s < t$ : compute the exit and adoption thresholds  $\beta_{t,s}$  and  $\gamma_{t,s}$  using Equations (IA40) and (IA41).
    - i. For each  $g, s < t$ , set  $T_{g,s}^{new} = (T_{g,s} \cap (\beta_{t,s}, \infty)) \cap (0, \gamma_{t,s})$ .
    - ii. Set  $T_{g,t}^{new} = (T_{g,s} \cap (\beta_{t,s}, \infty)) \cap (\gamma_{t,s}, \infty)$ .
    - iii. Set  $T_{t,t} = (\alpha_t, \infty)$  for  $\alpha_t$  defined in Equation (IA42).

---

<sup>37</sup>By the property of the power law,  $\alpha_g$  is a sufficient statistic to compute  $\Psi_t$ .

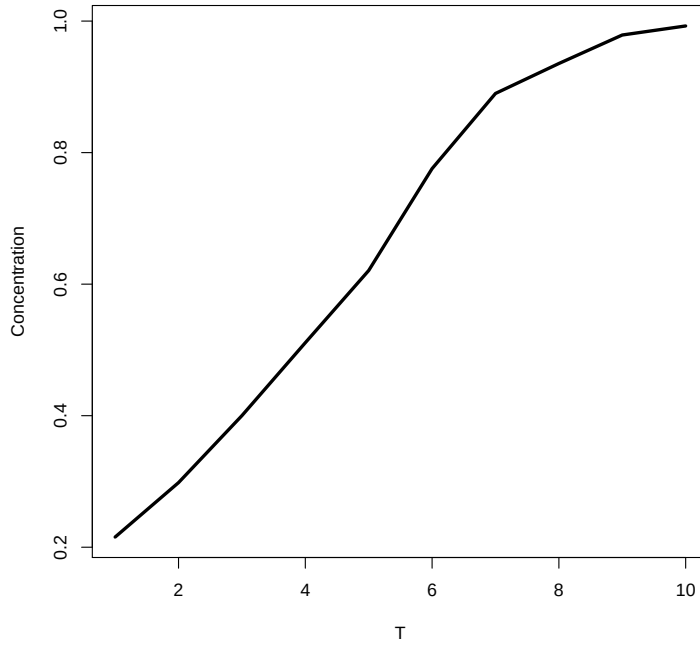


Figure IA24. Simulated Concentration Dynamics

*Notes:* This figure presents a numerical illustration of concentration (measured as top 1% sales share) over time in the dynamic model. Each period is set to be one decade.

- (c) For  $g < t$ : compute the total remaining number of (non-exiting) firms as a share of the total number of firms at time  $t - 1$ .
- (d) Verify that this exit share is equal to the target, and adjust proposed  $\Psi_t$ .
3. Finally, we compute the implied  $N_{t,t}$  using Equation (IA43).
4. Record the new  $T_{g,s}$  for  $g, s \leq t$ , the productivity thresholds  $\alpha_g$  for  $g = 1, \dots, t$ , and finally the mass of entering firms:  $N_1, N_2, \dots, N_t$ .

Figure IA24 provides a numerical illustration where we set each time period to be one decade (each time period corresponds to the introduction of the new generation of technology). We set  $h = 1.4$  to roughly correspond to the growth in aggregate output in a decade, and  $\mu = 0.2$  as the average markup. We set the idiosyncratic productivity threshold  $\alpha_t$  such that 30% of existing firms exit in each decade.<sup>38</sup> The figure shows the resulting concentration dynamics, as measured by the top 1% sales share. In the model, With the introduction of each generation of technology that has increasing stronger economies of scale, concentration rises over time.

<sup>38</sup>We set the remaining parameters to the following values:  $\bar{Y} = 1000, \phi = 3, \eta = 4, \sigma = 3, f = 2, k = 3$ .



### IA3.3 Trade Analyses

We now modify the basic framework in Appendix IA3.1 to analyze the impact of changes in trade barriers as well as possible interactions with economies of scale. For simplicity, we consider the symmetric case of two economies, “home” and “abroad” (given the symmetry of the two economies, we do not use additional notations to label each economy). The two economies can represent the U.S. and foreign countries for international trade, or the east coast and the west coast for domestic trade. We use “exports” to refer to goods sold outside of the local economy and “imports” to refer to goods purchased from outside. Since our data capture concentration in production activities in the U.S., when the two economies in the model are interpreted as the U.S. and abroad, we are interested in the share of top U.S. firms in total U.S. production; when the two economies are interpreted as east coast and west coast, we are interested in the share of top firms in both economies in the total production of the two economies. In the baseline analysis, we consider businesses’ sales including exports. When the two economies represent U.S. and abroad, we also examine “concentration excluding exports,” namely top U.S. businesses’ domestic sales over all U.S. businesses’ domestic sales.

For simplicity, each economy has the same aggregate demand and the same CES setup, where firms face demand  $\bar{Y} \cdot \left(\frac{p_i}{P}\right)^{-\sigma}$  with an exogenous markup  $\mu$  domestically. Following Melitz (2003), we assume that there is a standard iceberg cost  $\tau > 1$  (which raises the marginal cost of exports by a factor  $\tau$ ), as well as a fixed cost of exporting  $\phi_x$ . Following Melitz (2003), we assume that there is a standard iceberg cost  $\tau > 1$  (which raises the marginal cost by a factor of  $\tau$  for products shipped abroad), as well as a fixed cost of exporting  $\phi_x$ . We assume that exporting firms pass on the iceberg costs to consumers, i.e.  $p_{i,exp} = \tau p_i = \tau \frac{1+\mu}{a_i}$ . The key parameter for trade barrier is therefore  $\tau$ .

We start with the setting where all firms use the same constant-returns-to-scale technology, and only differ in their idiosyncratic productivity  $a_i$ . In other words, we shut down technologies with economies of scale and focus on trade barriers only. We obtain two key predictions. First, the top business share is hump-shaped in trade barriers  $\tau$ . For sufficiently high barriers to trade ( $\tau$  sufficiently large), lowering barriers to trade increases concentration: when  $\tau$  starts from a high level and falls, relatively more efficient firms export and top businesses’ shares increase. However, when  $\tau$  is already low, then further decreases in  $\tau$  will primarily make mediocre firms export (top firms already export); in this case, concentration can decrease rather than increase. Second, when the two economies are interpreted as the U.S. and abroad, then U.S. “concentration excluding exports” remains unchanged. Accordingly, if changes in international trade barriers represent the only force, then we may expect U.S. “concentration excluding exports” to be stable.

We then analyze a setting where trade barriers coexist with the technology with economies of scale (i.e., the new technology has  $h > 1$ ). In this case, when trade barriers fall (smaller  $\tau$ ), a higher fraction of firms will adopt the new technology with greater scalability, and concentration can increase. In other words, having broader markets can amplify economies of scale. In addition, in this case even “concentration excluding exports” can increase.

#### IA3.3.1 Homogeneous Technology

We start with homogeneous technology: all firms use the same constant-returns-to-scale technology but differ in their idiosyncratic productivity  $a_i$  (which corresponds to the case of  $h = 1$  in the baseline model). For simplicity, we also assume that firms are only active for one period: the conclusions of our

model remain unchanged in the standard specification.<sup>39</sup> Then, the per-period profits are reduced to a single period profit:  $\Pi(a_i) = \pi(a_i) - \phi$ .

Given the aggregate price index  $P$ , total profits of operating firms that do not export are given by:

$$\pi_{\text{no export}}(a_i) = \bar{Y} P^\sigma \left\{ \left( \frac{1 + \mu}{a_i} \right)^{-\sigma} \cdot \frac{\mu}{a_i} \right\} - \phi, \quad (\text{IA47})$$

whereas total profits of operating firms that export are given by:

$$\pi_{\text{export}}(a_i) = \bar{Y} P^\sigma \left\{ \left( \frac{1 + \mu}{a_i} \right)^{-\sigma} \cdot \frac{\mu}{a_i} + \underbrace{\left( \frac{\tau(1 + \mu)}{a_i} \right)^{-\sigma} \cdot \frac{\tau\mu}{a_i}}_{\text{export profits}} \right\} - \phi - \phi_x. \quad (\text{IA48})$$

From the operating constraint and the exporting constraint, one can implicitly define two thresholds, given by:

$$\begin{aligned} \bar{Y} P^\sigma \left( \frac{1 + \mu}{a^*} \right)^{-\sigma} \frac{\mu}{a^*} &= \phi \\ \bar{Y} P^\sigma \left( \frac{1 + \mu}{a_{\text{exp}}} \right)^{-\sigma} \frac{\mu}{a_{\text{exp}}} \cdot \tau^{1-\sigma} &= \phi_x, \end{aligned} \quad (\text{IA49})$$

where  $a^*$  is the minimum productivity threshold required to operate profitably, and  $a_{\text{exp}}$  is the minimum productivity required for exporting to be profitable. Combining the two equations yields:

$$\frac{a_{\text{exp}}}{a^*} = \tau \left( \frac{\phi_x}{\phi} \right)^{\frac{1}{\sigma-1}} \quad (\text{IA50})$$

Intuitively, fewer firms export ( $\frac{a_{\text{exp}}}{a^*}$  is high) when  $\tau$  is high (greater iceberg costs) or  $\phi_x$  (fixed cost of exporting) is high. We shall assume (for convenience) that either cost is sufficiently high such that  $a_{\text{exp}} \geq a^*$ .

Given the relative cutoffs  $a^*$  and  $a_{\text{exp}}$ , there are three types of firms.

1. Firms that export: their sales are given by  $\bar{Y} P^\sigma (1 + \mu)^{1-\sigma} \cdot a_i^{\sigma-1} (1 + \tau^{1-\sigma})$ .
2. Firms that do not export: their sales are given by  $\bar{Y} P^\sigma (1 + \mu)^{1-\sigma} \cdot a_i^{\sigma-1}$ .
3. Imports from firms abroad: their sales (in the domestic market) are given by  $\tau^{1-\sigma} \bar{Y} P^\sigma (1 + \mu)^{1-\sigma} \cdot a_i^{\sigma-1}$ .

In other words, among active firms, exporters have revenues proportional to  $(1 + \tau^{1-\sigma}) a_i^{\sigma-1}$ , where  $a_i \geq a_{\text{exp}}$ , non-exporters have revenues proportional to  $a_i^{\sigma-1}$ , where  $a_{\text{exp}} \geq a_i \geq a^*$ , and importers have revenues proportional to  $\tau^{1-\sigma} a_i^{\sigma-1}$ , where  $a_i \geq a_{\text{exp}}$ . We maintain our assumption that the idiosyncratic productivity is Pareto-distributed with  $P(a \geq x) \propto x^{-k}$ , with  $k \geq \sigma - 1$ .

---

<sup>39</sup>The only difference this additional assumption introduces is the addition of a term involving the discount rate  $R$ , with all of the conclusions remaining unchanged.

Following our empirical exercise, we measure concentration in production activities. In other words, we consider total production by domestic firms *including* exports made by domestic exporters and excluding imports from foreign firms. Assume for simplicity that the cutoff for the concentration ratio is such that the non-exporting firms are not part of the top 1%. In this case, it suffices (given that the most productive firms export) to consider the sales (including exports) of the firms for which  $a \geq a^* \cdot (0.01)^{-\frac{1}{k}}$ . Then, the expressions for the top 1% sales share with and without exports are given by the following respectively:

$$\begin{aligned}\zeta_{1\%, \text{with exports}} &= \frac{(1 + \tau^{1-\sigma}) \cdot (0.01)^{1-\frac{\sigma-1}{k}}}{1 + \tau^{1-\sigma} \cdot \left(\frac{a_{exp}}{a^*}\right)^{-(k-(\sigma-1)T)}} = \frac{(1 + \tau^{1-\sigma}) \cdot (0.01)^{1-\frac{\sigma-1}{k}}}{1 + \tau^{-k} \left(\frac{\phi_x}{\phi}\right)^{1-\frac{k}{\sigma-1}}} \\ \zeta_{1\%, \text{without exports}} &= (0.01)^{1-\frac{\sigma-1}{k}}\end{aligned}\tag{IA51}$$

One can then derive the following comparative static of the top 1% sales share with respect to trade barrier  $\tau$ .

**Proposition IA6.** *The top share is hump-shaped in  $\tau$ : for sufficiently high barriers to trade ( $\tau$  sufficiently high), lowering barriers to trade increases concentration. Furthermore, “concentration excluding exports” remains the same as barriers to trade change.*

### IA3.3.2 Incorporating Scalable Technology

Next, we augment the basic trade model and include the two types of technologies specified in our main model. In other words, firms with idiosyncratic productivity  $a_i$  can choose to invest in the standard technology above or alternatively in a new technology that requires  $\phi \cdot h^\eta > \phi$  of spending but provides per-unit productivity  $a_i \cdot h > a_i$ .

For ease of analysis, we assume that technological innovations that affect the scalability of the new technology (the parameter  $h$ ) is shared across both markets. In this case, there are two dimensions of firm optimization: 1) whether to export or not, and 2) whether to adopt the new technology. There are three possible equilibria:

1. The least efficient firms neither export nor adopt the new technology. Mediocre firms export but with old technology. Efficient firms export with new technology.
2. The least efficient firms neither export nor adopt the new technology. Mediocre firms adopt the new technology but do not export. Efficient firms export with new technology.
3. The least efficient firms neither export nor adopt the new technology. The rest export with new technology.

Which of the three possibilities ends up occurring depends on the relative cost of the new technology (indexed by  $\eta$ ) and the cost of export ( $\phi_x$ ). For convenience, we shall assume the following regarding the relative values of  $\eta$  and  $\phi_x$ , which guarantees the third possibility, the simplest case to analyze. Nonetheless, the general conclusions and comparative statics remain largely unchanged if we relax the assumption.

**Assumption IA4.** Let  $h, \eta, \phi_x$  and  $\tau$  satisfy the following:

$$\frac{h^{\sigma-1} - 1}{h^\eta - 1} \cdot (1 + \tau^{\sigma-1}) \geq \frac{\phi}{\phi_x} \geq \frac{h^{\sigma-1} - 1}{h^\eta - 1} \tau^{1-\sigma} h^{1-\sigma} \quad (\text{IA52})$$

**Proposition IA7.** Under the above assumption, there are two types of firms in the economy: firms with idiosyncratic productivity  $a^{**} \geq a \geq a^*$  that neither export nor adopt the new technology, and firms with idiosyncratic productivity  $a \geq a^{**}$  who do both. The two thresholds satisfy:

$$\frac{a^{**}}{a^*} = \left( \frac{(h^\eta - 1) + \frac{\phi_x}{\phi}}{(1 + \tau^{1-\sigma})h^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}}. \quad (\text{IA53})$$

First, in the presence of heterogeneous technology and economies of scale, changes in the barriers to trade now has an effect on both standard top sales concentration, as well as the export-excluded top sales concentration.

**Corollary IA1.** Consider a reduction in barriers to trade ( $\tau$ ), then more firms (as a share of total firms) will adopt the new technology:  $\frac{a^{**}}{a^*}$  falls. Assuming as before that the new technology is sufficiently marginal (i.e.  $h = 1 + \nu$ ) and barriers to trade ( $\tau$ ) sufficiently high, then the top business share will increase.

**Corollary IA2.** Sales concentration excluding exports decreases with  $\tau$  and increases with  $h$ .

Contrary to the case with homogeneous technology (Proposition IA6), here lower trade barriers can increase concentration even when one excludes exports. This is because lower trade barriers lead to greater adoption of the scalable technology. Meanwhile, holding trade barriers fixed, the increase in economies of scale (higher  $h$ ) raises concentration (as before) as well as “concentration excluding exports.”

**Corollary IA3.** Consider a marginal increase in the scalability of the new technology:  $h = 1 + \nu$  with  $\nu \mapsto 0$ . The top business share will increase, along with total exports and imports as a share of gross output.

Corollary IA1 and Corollary IA3 illustrate that trade and the scalable technology are complements. Lower barriers to trade encourage firms to adopt the scalable technology, while enhancement in the scalable technology increases the ability of efficient firms to reach foreign markets and increase the trade volume.

The above analyses are framed primarily as international trade (“home” and “abroad,” “exports” and “imports”). The same framework can be used to think about the integration of domestic markets. In this case, we can think of “home” and “abroad” as say “east coast” and “west coast.” Concentration in the U.S. would be the sales share of top businesses in these two economies in the total sales of these two economies. When the two economies are symmetric, this is the same as the top business share in each economy, so the expressions for concentration are the same as before. The only difference for the case of domestic trade is that we cannot easily measure “exports” and “imports” (i.e., trade across different U.S. regions), and we cannot obtain “concentration excluding exports” in the data.

### IA3.3.3 Proofs

*Proof of Proposition IA6.* Using standard power-law identities,<sup>40</sup> we obtain that the sales share of the top 1% is given by:

$$\frac{(1 + \tau^{1-\sigma}) \cdot (0.01)^{1-\frac{\sigma-1}{k}}}{1 + \tau^{1-\sigma} \cdot \left(\frac{a_{exp}}{a^*}\right)^{-(k-(\sigma-1)T)}} = \frac{(1 + \tau^{1-\sigma}) \cdot (0.01)^{1-\frac{\sigma-1}{k}}}{1 + \tau^{-k} \left(\frac{\phi_x}{\phi}\right)^{1-\frac{k}{\sigma-1}}}. \quad (\text{IA54})$$

Taking the comparative static with respect to  $\tau$ , we have that the share is increasing in  $\tau$  iff:

$$\begin{aligned} 0 &< k \cdot \frac{\left(\frac{\phi_x}{\phi}\right)^{1-\frac{k}{\sigma-1}} \tau^{-k-1}}{1 + \tau^{-k} \left(\frac{\phi_x}{\phi}\right)^{1-\frac{k}{\sigma-1}}} - \frac{(\sigma-1)\tau^{-\sigma}}{1 + \tau^{1-\sigma}} \\ &\iff \left(1 + \tau^{-k} \left(\frac{\phi_x}{\phi}\right)^{1-\frac{k}{\sigma-1}}\right) (\sigma-1)\tau^{-\sigma} < \left(k \cdot \left(\frac{\phi_x}{\phi}\right)^{1-\frac{k}{\sigma-1}} \tau^{-k-1}\right) (1 + \tau^{1-\sigma}) \quad (\text{IA55}) \\ &\iff (\sigma-1) \left(\frac{\phi_x}{\phi}\right)^{\frac{k}{\sigma-1}-1} < k \cdot \tau^{-(k-(\sigma-1))} + (k - (\sigma-1)) \cdot \tau^{-k}. \end{aligned}$$

Note that the left hand side of the last expression is fixed in  $\tau$ , while the right hand side is a strictly decreasing function in  $\tau$  that eventually converges to 0. Furthermore, at the minimum value of  $\tau = \left(\frac{\phi_x}{\phi}\right)^{-\frac{1}{\sigma-1}}$ , one has that the right hand side evaluates to:  $k \cdot \left(\frac{\phi_x}{\phi}\right)^{\frac{k}{\sigma-1}-1} + (k - (\sigma-1)) \cdot \left(\frac{\phi_x}{\phi}\right)^{\frac{k}{\sigma-1}}$ , which is greater than the left hand side. Thus, by the single-crossing condition, we have that the original top share is hump-shaped in  $\tau$ . □

*Proof of Proposition IA7.* First, we consider firms that are on the cutoff of exporting:

$$\bar{Y} P^\sigma \left(\frac{1+\mu}{a}\right)^{-\sigma} \frac{\mu}{a} \cdot \tau^{1-\sigma} = \phi_x. \quad (\text{IA56})$$

These firms then adopt the new technology iff the extra revenue justifies the higher cost of the technology:

$$(h^{\sigma-1} - 1) \cdot (1 + \tau^{1-\sigma}) \cdot \bar{Y} P^\sigma \left(\frac{1+\mu}{a}\right)^{-\sigma} \frac{\mu}{a} = (h^{\sigma-1} - 1) \cdot (1 + \tau^{\sigma-1}) \phi_x \geq (h^\eta - 1) \phi. \quad (\text{IA57})$$

In other words, we would like to assume:

$$\frac{h^{\sigma-1} - 1}{h^\eta - 1} \cdot (1 + \tau^{\sigma-1}) \phi_x \geq \phi. \quad (\text{IA58})$$

---

<sup>40</sup>Specifically,  $\int_\alpha^\infty k(a^*)^k a^{-(k+1)} \cdot a^{\sigma-1} da = \frac{k}{k+1-\sigma} (a^*)^{\sigma-1} \left(\frac{\alpha}{a^*}\right)^{-k+\sigma-1}$ .

Second, we consider firms that are on the cutoff of adopting the new technology:

$$(h^{\sigma-1} - 1) \bar{Y} P^\sigma \left( \frac{1 + \mu}{a} \right)^{-\sigma} \frac{\mu}{a} = (h^\eta - 1) \phi. \quad (\text{IA59})$$

For these firms to export, the increase in revenue is given by:

$$\tau^{\sigma-1} \cdot h^{\sigma-1} \bar{Y} P^\sigma \left( \frac{1 + \mu}{a} \right)^{-\sigma} \frac{\mu}{a}, \quad (\text{IA60})$$

which must be balanced by the extra cost  $\phi_x$ .

Thus, for the marginal adopters of the new technology to export, we must have:

$$\tau^{\sigma-1} \cdot h^{\sigma-1} \bar{Y} P^\sigma \left( \frac{1 + \mu}{a} \right)^{-\sigma} \frac{\mu}{a} = \tau^{\sigma-1} \cdot h^{\sigma-1} \cdot \frac{h^\eta - 1}{h^{\sigma-1} - 1} \phi \geq \phi_x, \quad (\text{IA61})$$

or in other words:

$$\phi \geq \frac{h^{\sigma-1} - 1}{h^\eta - 1} \tau^{1-\sigma} h^{1-\sigma} \phi_x. \quad (\text{IA62})$$

□

*Proof of Corollary IA3.* The top 1% sales share is given by:

$$\frac{(1 + \tau^{1-\sigma}) h^{\sigma-1} \cdot (0.01)^{1 - \frac{\sigma-1}{k}}}{1 + ((1 + \tau^{1-\sigma}) h^{\sigma-1} - 1) \cdot \left( \frac{a^{**}}{a^*} \right)^{-(k-(\sigma-1))}} = \frac{(1 + \tau^{1-\sigma}) h^{\sigma-1} \cdot (0.01)^{1 - \frac{\sigma-1}{k}}}{1 + ((1 + \tau^{1-\sigma}) h^{\sigma-1} - 1)^{\frac{k}{\sigma-1}} \cdot \left( (h^\eta - 1) + \frac{\phi_x}{\phi} \right)^{1 - \frac{k}{\sigma-1}}}, \quad (\text{IA63})$$

and the total export share of this economy is given by:

$$\frac{\tau^{1-\sigma} h^{\sigma-1} \cdot \left( \frac{(h^\eta - 1) + \frac{\phi_x}{\phi}}{(1 + \tau^{1-\sigma}) h^{\sigma-1} - 1} \right)^{1 - \frac{k}{\sigma-1}}}{1 + ((1 + \tau^{1-\sigma}) h^{\sigma-1} - 1)^{\frac{k}{\sigma-1}} \cdot \left( (h^\eta - 1) + \frac{\phi_x}{\phi} \right)^{1 - \frac{k}{\sigma-1}}} = \frac{\tau^{1-\sigma} h^{\sigma-1} \cdot \left( \frac{a^{**}}{a^*} \right)^{-k+(\sigma-1)}}{1 + ((1 + \tau^{1-\sigma}) h^{\sigma-1} - 1) \cdot \left( \frac{a^{**}}{a^*} \right)^{-k+(\sigma-1)}}. \quad (\text{IA64})$$

We consider the impact of a marginal increase in  $h$  from 1 to  $1 + \epsilon$ . We have:  $\frac{a^{**}}{a^*} = \left( \frac{(h^\eta - 1) + \frac{\phi_x}{\phi}}{(1 + \tau^{1-\sigma}) h^{\sigma-1} - 1} \right)^{\frac{1}{\sigma-1}} \approx \left( \frac{\phi_x}{\phi} \right)^{\frac{1}{\sigma-1}} \tau$ , which is independent of  $h$ .

Then, both the top business share and the export share of output are approximately proportional to a rational function in  $h^{\sigma-1}$  of the form:

$$\frac{h^{\sigma-1}}{A \cdot h^{\sigma-1} + B}, \quad (\text{IA65})$$

where  $A = (1 + \tau^{1-\sigma}) \left( \frac{a^{**}}{a^*} \right)^{-k+(\sigma-1)}$ ,  $B = 1 - \left( \frac{a^{**}}{a^*} \right)^{-k+(\sigma-1)}$ . As  $a^{**} \geq a^*$  and  $k > \sigma - 1$ ,  $A, B > 0$ , which implies that the above function is increasing in  $h$ .

Finally, the imports from foreign firms are equal (via symmetry) to the net exports of exporting firms (as a share of GDP).

□

*Proof of Corollary IA1.* It suffices to show that  $\frac{(h^\eta - 1) + \frac{\phi_x}{\phi}}{(1 + \tau^{1-\sigma})h^{\sigma-1} - 1}$  is increasing in  $\tau$ . This follows trivially from the observation that as  $\sigma > 1$ , the denominator is decreasing in  $\tau$  and the original expression is thus increasing in  $\tau$ .

The expression for concentration is given by the following:

$$\frac{(1 + \tau^{1-\sigma})h^{\sigma-1} \cdot (0.01)^{1-\frac{\sigma-1}{k}}}{1 + ((1 + \tau^{1-\sigma})h^{\sigma-1} - 1)^{\frac{k}{\sigma-1}} \cdot \left((h^\eta - 1) + \frac{\phi_x}{\phi}\right)^{1-\frac{k}{\sigma-1}}}. \quad (\text{IA66})$$

To see the comparative static of the above expression as a function of  $\tau$ , it is convenient to denote  $A(\tau) = 1 + \tau^{1-\sigma}$ , which is a decreasing function of  $\tau$ . Factoring out terms that are independent of  $\tau$ , we see that it suffices to determine the comparative static of

$$\frac{A(\tau)}{1 + (A(\tau)h^{\sigma-1} - 1)^{\frac{k}{\sigma-1}} \cdot B}, \quad (\text{IA67})$$

with respect to  $A(\tau)$ . Taking the log-derivative of the above expression with respect to  $A$  yields:

$$\begin{aligned} & (1 + (Ah^{\sigma-1} - 1)^{\frac{k}{\sigma-1}} \cdot B) - A \cdot \frac{k}{\sigma-1} \cdot h^{\sigma-1} \cdot (Ah^{\sigma-1} - 1)^{\frac{k}{\sigma-1}-1} \cdot B \\ &= 1 + (Ah^{\sigma-1} - 1)^{\frac{k}{\sigma-1}-1} \cdot B \cdot \left( (Ah^{\sigma-1} - 1) - A \frac{k}{\sigma-1} h^{\sigma-1} \right). \end{aligned} \quad (\text{IA68})$$

Note that as  $\tau \mapsto \infty$  and  $h \mapsto 1$ , the above term converges to  $1 > 0$ . Consequently, the above expression is increasing in  $A$  as  $h \mapsto 1$  and  $\tau$  is sufficiently large. As  $A(\tau)$  is a decreasing function in  $\tau$ , this means that the original expression for concentration, under the above assumptions, is decreasing in  $\tau$ , as desired.

□

*Proof of Corollary IA2.* When we exclude exports from the original expression, we have the following expression for “concentration excluding exports”:

$$\frac{h^{\sigma-1} \cdot (0.01)^{1-\frac{\sigma-1}{k}}}{1 + (h^{\sigma-1} - 1) \cdot \left(\frac{a^{**}}{a^*}\right)^{-(k-(\sigma-1))}}. \quad (\text{IA69})$$

Naturally, since  $\tau$  does not enter the expression except indirectly through  $\frac{a^{**}}{a^*}$ , which is increasing in  $\tau$ , we obtain that the top business share is also increasing in  $\tau$ . The comparative statics with respect to  $h$  follow the same logic as the case without trade.

□