Internet Appendix:
100 Years of Rising Corporate Concentration

Spencer Y. Kwon Yueran Ma Kaspar Zimmermann*

May 2024

This Internet Appendix has two sections. Section IA1 presents additional results. Section IA2 provides more information about data and variable construction: Section IA2.1 explains the SOI data in detail; Section IA2.2 explains additional historical data on the top 100 asset share in manufacturing; Section IA2.3 explains the construction of the main variables in our analyses; Section IA2.4 explains the lognormal interpolation of top shares for robustness checks.

*Kwon: Brown University (spencer_kwon@brown.edu); Ma: University of Chicago Booth School of Business and NBER (yueran.ma@chicagobooth.edu); Zimmermann: Frankfurt School of Finance and Management (kaspar.zimmermann@fs.de).

IA1 Additional Results

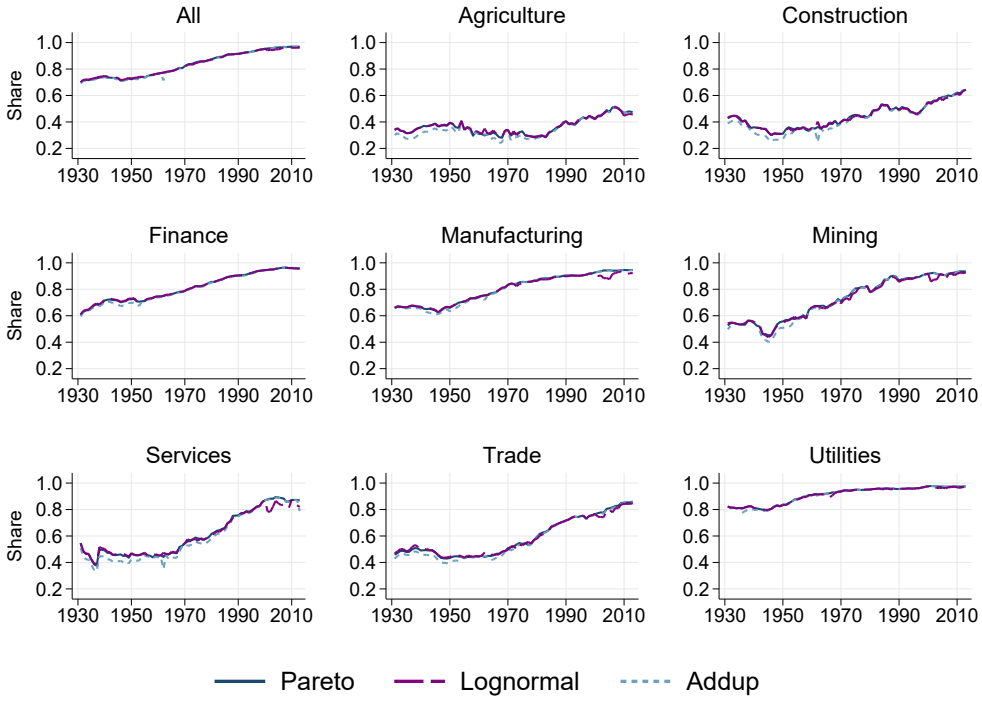


Figure IA1. Comparison of Different Methods for Estimating Top Shares

Notes: This figure shows the share of the top 1% corporations by assets in total corporate assets, using the three methods explained in Section I. 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.

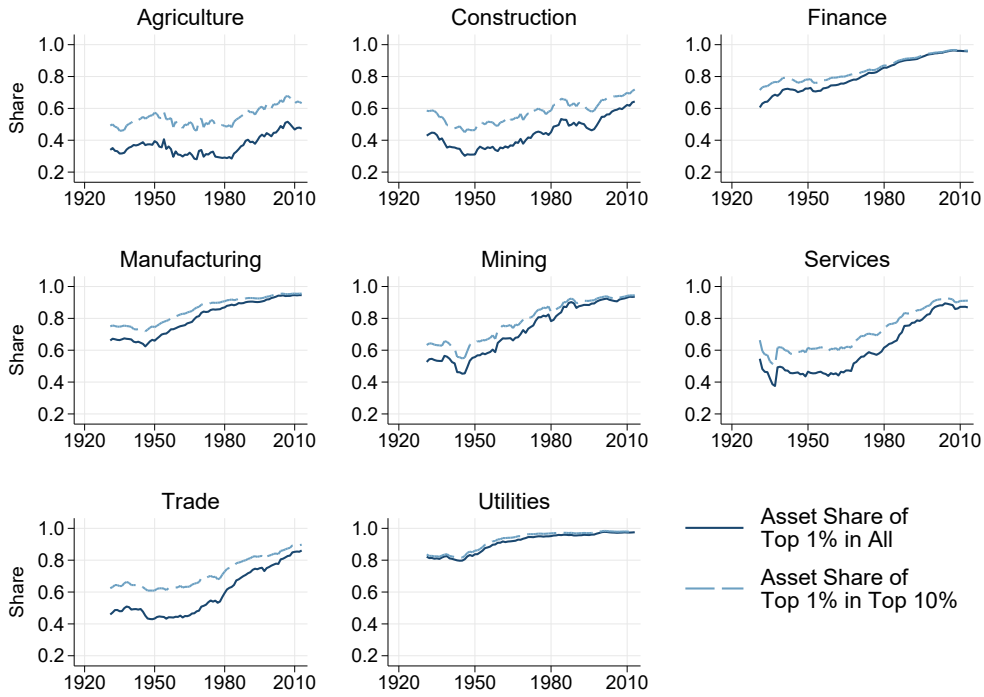
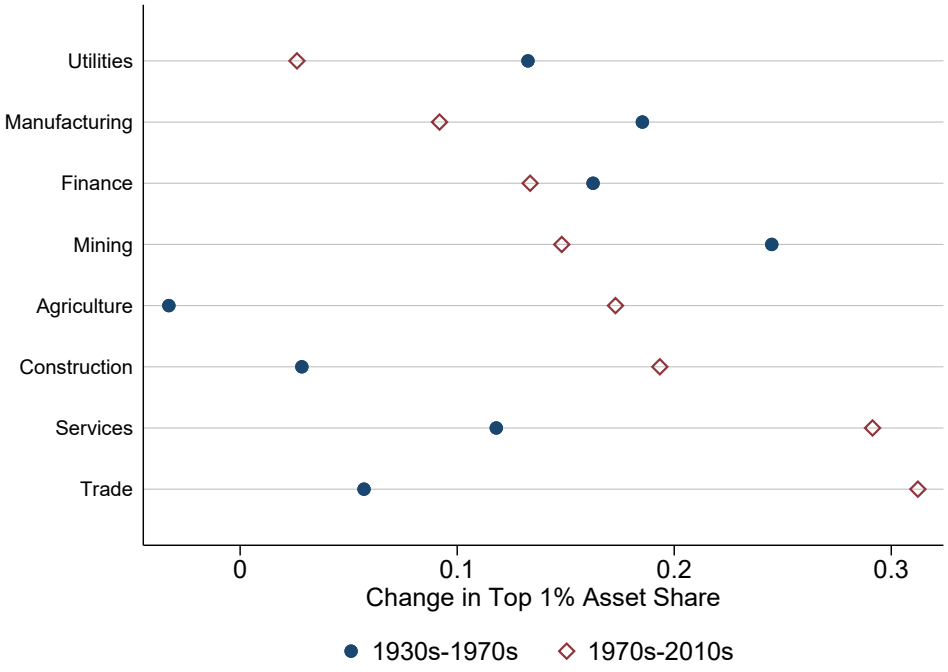


Figure IA2. Relative Shares among Top Corporations: Main Sectors by Assets

Notes: This figure shows the share of the top 1% corporations among the top 10% corporations in the main sectors. The solid line repeats the share of the top 1% corporations by assets in total corporate assets shown in Figure 3, and the dashed line shows the share of the top 1% corporations by assets in the top 10% corporations by assets.

Panel A. Main Sectors

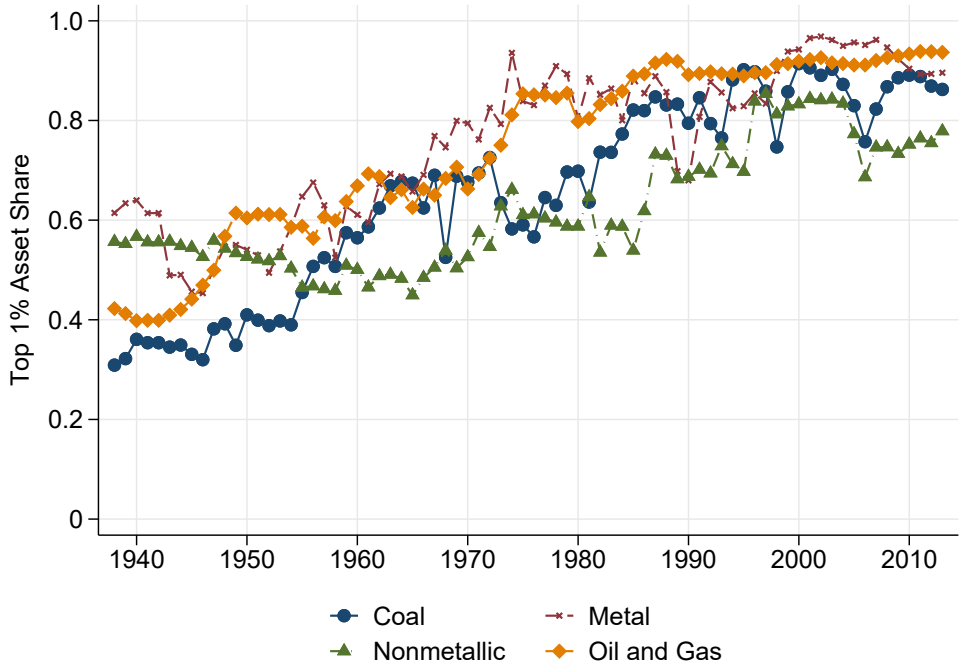


Panel B. Subsectors

Figure IA3. Rising Concentration in Earlier and Later Decades

Notes: This figure shows the share of the top 1% corporations by assets in total corporate assets in the 2010s relative to that in the 1970s (hollow red diamonds), and the share in the 1970s relative to that in the 1930s (solid blue circles), for main sectors in Panel A and subsectors in Panel B. The industries are sorted by the change between the 1970s and the 2010s.

Panel A. Mining



Panel B. Retail

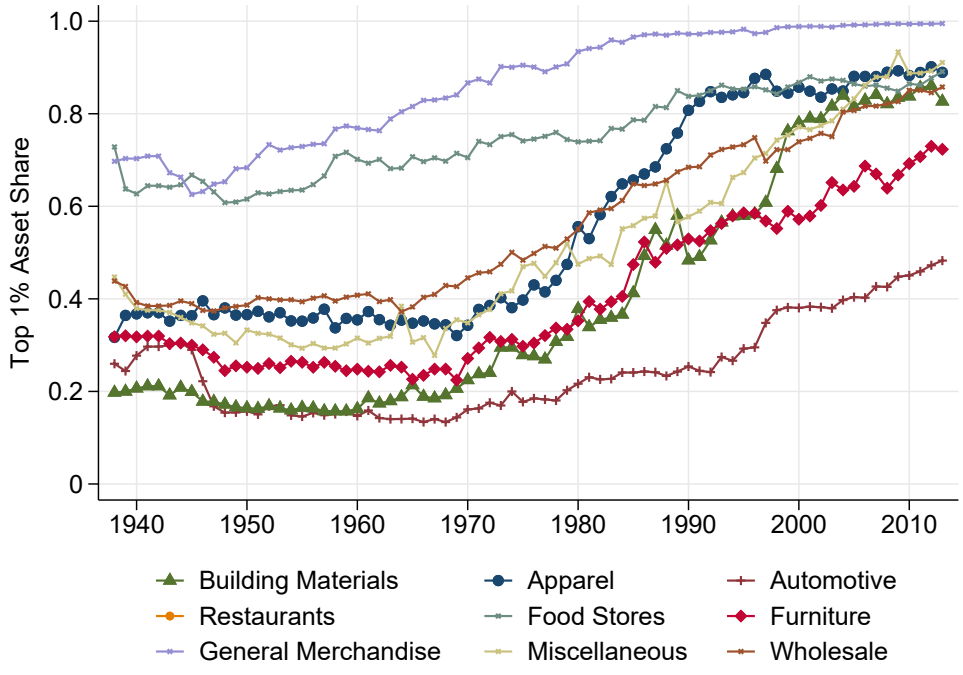


Figure IA4. Concentration within Subsectors

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

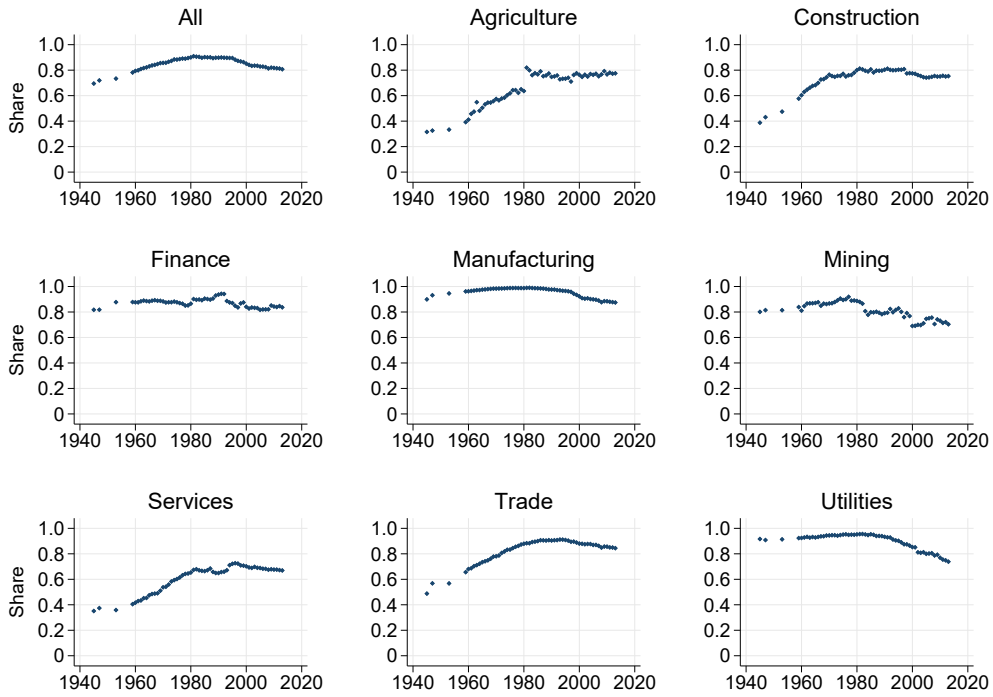
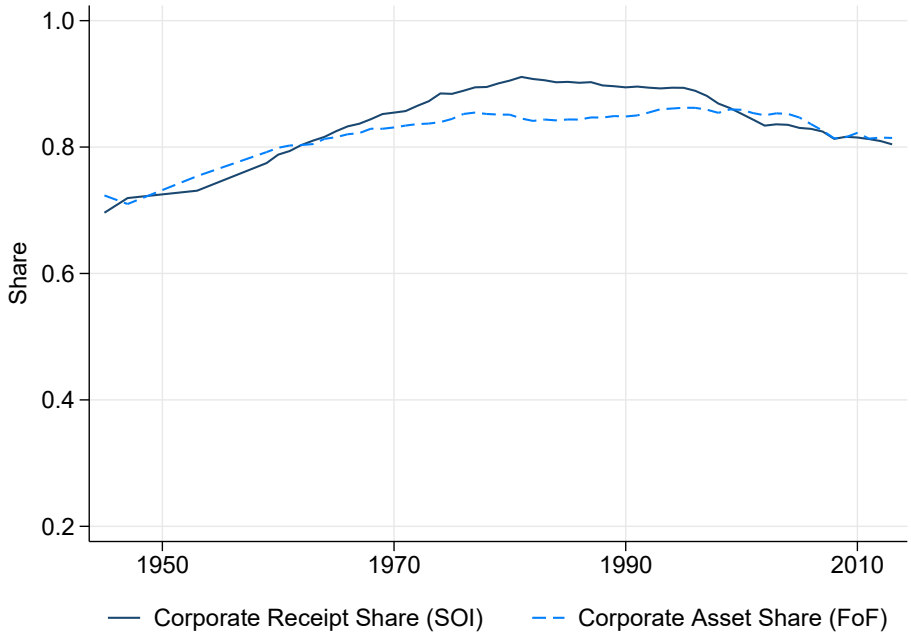


Figure IA5. Corporate Share in Receipts

Notes: This figure shows the share of corporations in the receipts of corporations, partnerships, and nonfarm proprietorships, for the aggregate and the main sectors.

Panel A. Nonfinancial Industries: Receipt Share and Asset Share



Panel B. Manufacturing: Receipt Share and Value Added Share

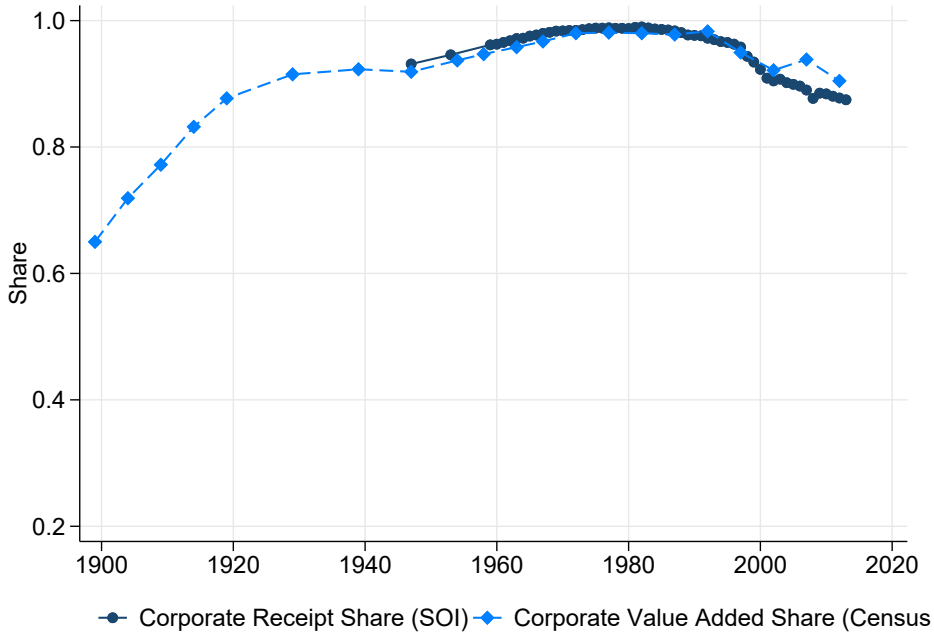


Figure IA6. Corporate Share Measures

Notes: In Panel A, the solid line shows the share of corporate receipts in corporate and noncorporate receipts for nonfinancial industries using SOI data. The dashed line shows the share of corporate assets in corporate and noncorporate assets for nonfinancial industries using Flow of Funds data (we exclude farm assets and residential real estate in Flow of Funds data to better match SOI data). In Panel B, the solid line with circles shows the share of corporate receipts in corporate and noncorporate receipts for manufacturing using SOI data. The dashed line with diamonds shows the share of corporate value added in corporate and noncorporate value added for manufacturing using manufacturing census data.

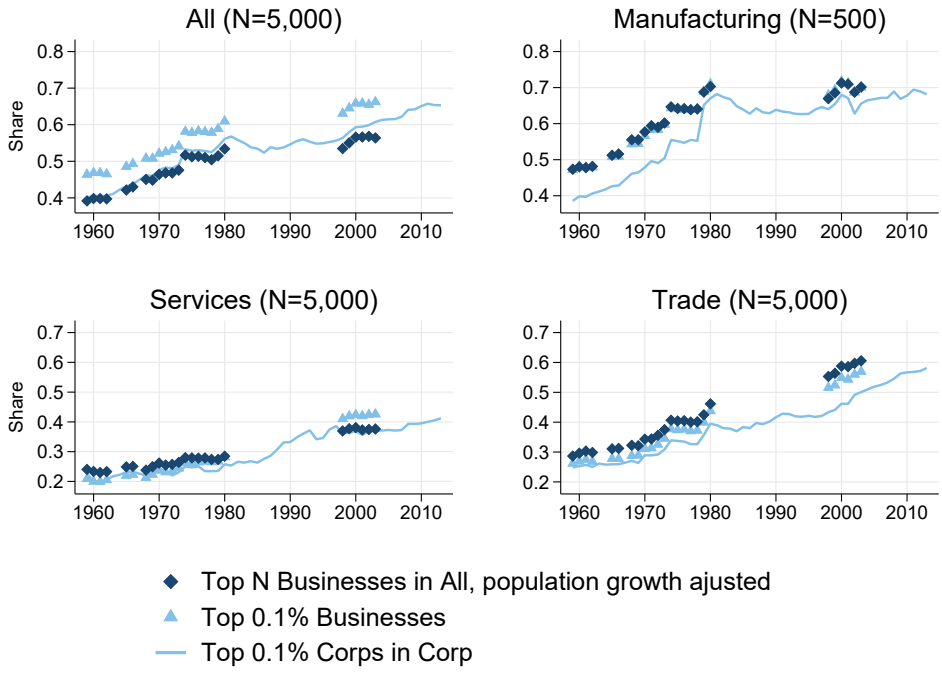


Figure IA7. Top N Receipt Shares with Population Growth Adjustment

Notes: This figure is the same as Panel A of Figure 7 except that the N for the dark blue diamonds grows at the rate of population. The value N is equal to 500 for manufacturing, and 5,000 for the aggregate, services, and retail and wholesale trade for 1980, and adjusted using the cumulative population growth rate relative to 1980 for other years. See Internet Appendix IA2.3 for details about variable construction.

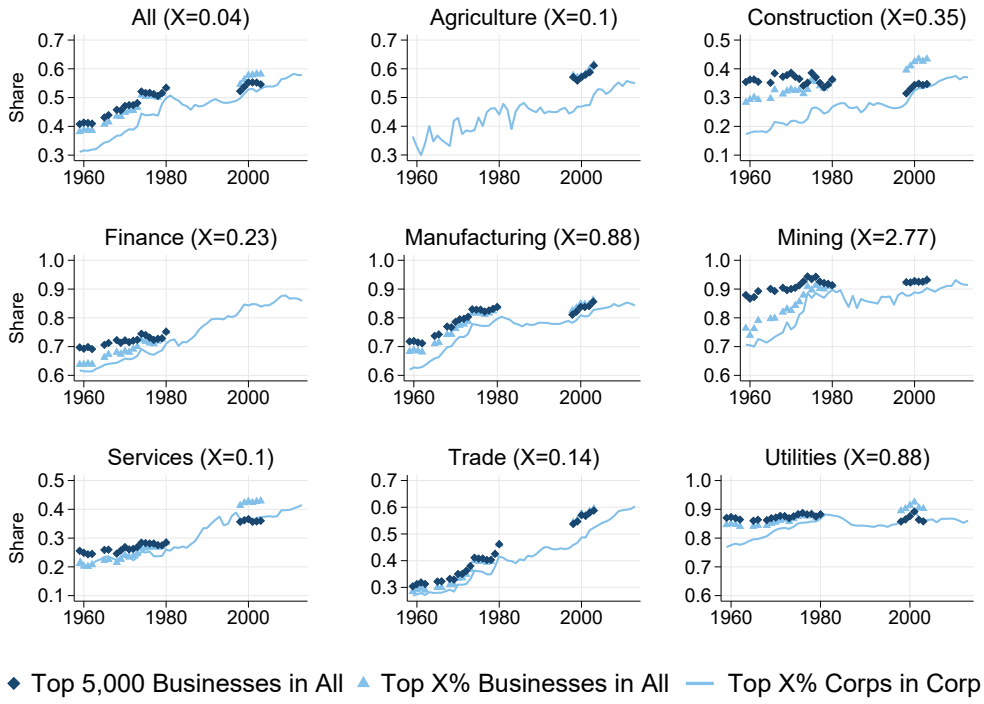


Figure IA8. Top 5,000 Receipt Shares

Notes: This figure shows the receipt share of the largest 5,000 businesses by receipts. The dark blue diamonds show the receipt share of the largest 5,000 businesses (corporations and noncorporations) by receipts among the receipts of all businesses. This is possible for years where we have data on the receipt size distribution of both corporations and noncorporations. For comparison, the light blue triangles show the receipt share of the top x% businesses by receipts among all businesses. The light blue line shows the receipt share of the top x% corporations by receipts among all corporations. The value of x% is computed as 5,000 divided by the total number of corporations and noncorporations in 1980, and x is labeled next to the sector name. See Internet Appendix IA2.3 for details about variable construction.

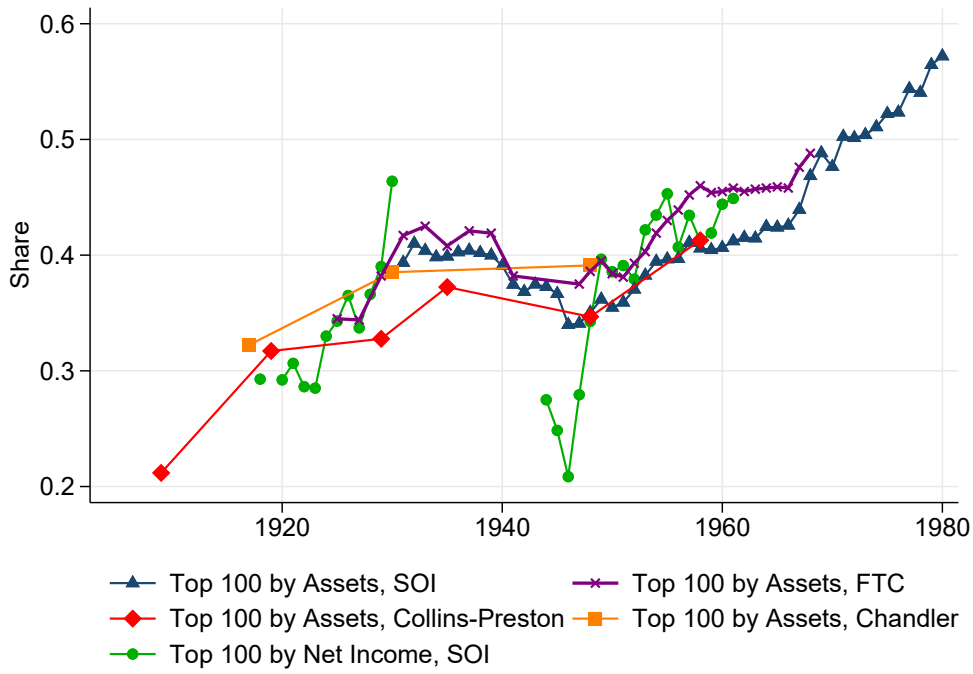
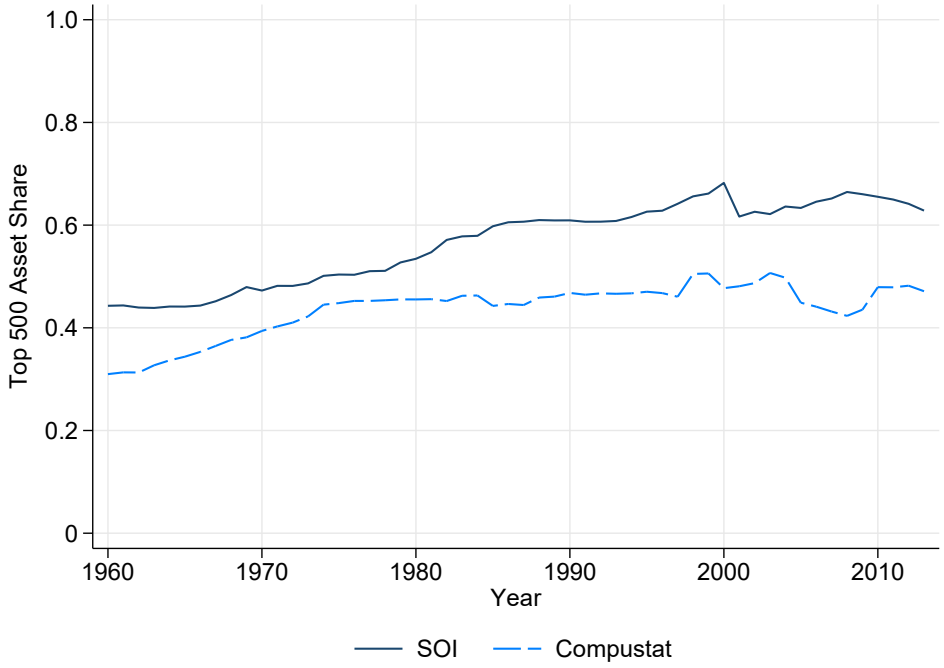


Figure IA9. Top 100 Asset Share in Manufacturing

Notes: This figure shows the asset share of the largest 100 manufacturers by assets estimated from multiple historical sources. The dark blue triangles show the top 100 asset share estimated using SOI data. The purple crosses show the top 100 asset share estimated by Federal Trade Commission (1969). The red diamonds show the assets of the largest 100 industrials listed by Collins and Preston (1961) relative to total assets estimated from SOI data. The orange squares show the assets of the largest 100 manufacturers from Chandler (1994) relative to total assets estimated from SOI data. The green circles show the top 100 net income share estimated using SOI data (restricted to businesses with positive net income). For total business assets in the denominator, we start with the assets of corporations and estimate the assets of noncorporations. We approximate the ratio of corporate to noncorporate assets in the early decades using the ratio of corporate to noncorporate value added from historical manufacturing census data (linearly interpolated between census years), as shown in Panel B of Figure IA6. See Internet Appendix IA2.2 for more details.

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



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

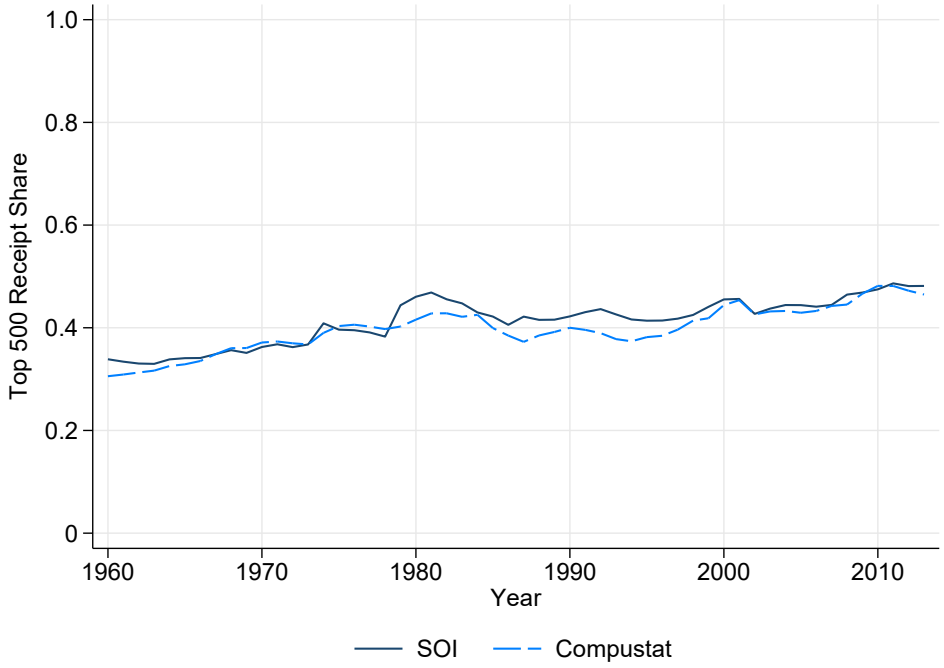


Figure IA10. Comparison with Compustat

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.

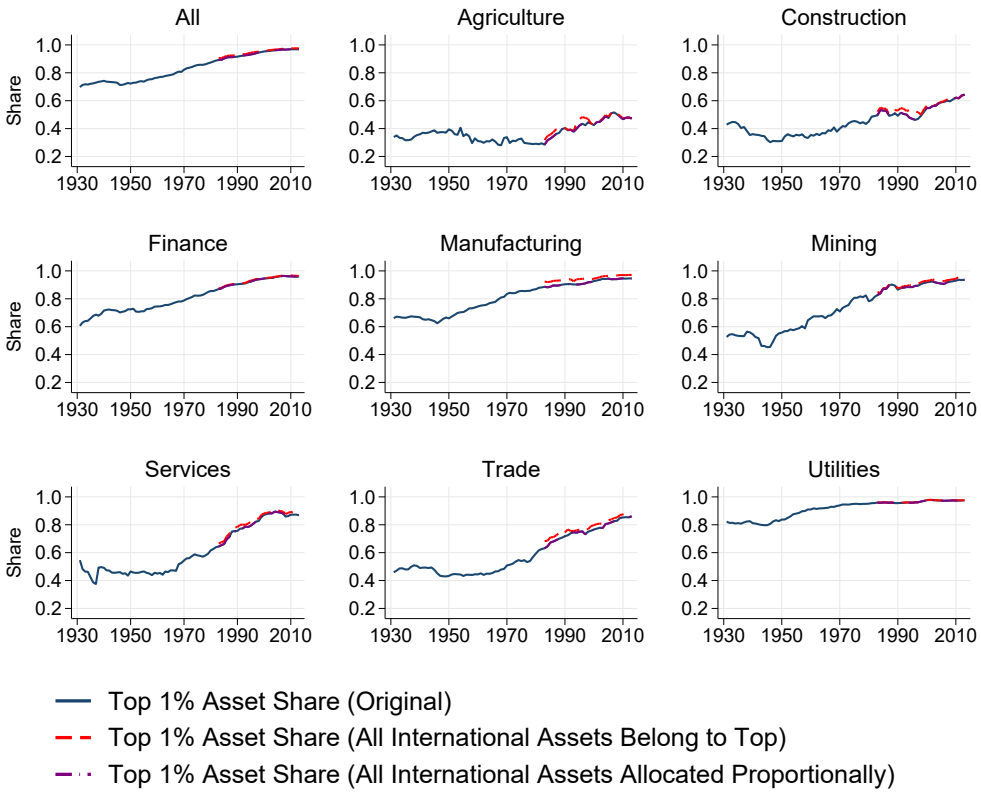


Figure IA11. Including International Assets

Notes: This figure shows the share of the top 1% corporations by assets in total corporate assets 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% corporations. The dash-dotted line shows the top 1% asset share when international assets are assigned to the top 1% and the bottom 99% corporations according to their domestic asset shares (using SOI data). See Internet Appendix IA2.3 for details about variable construction.

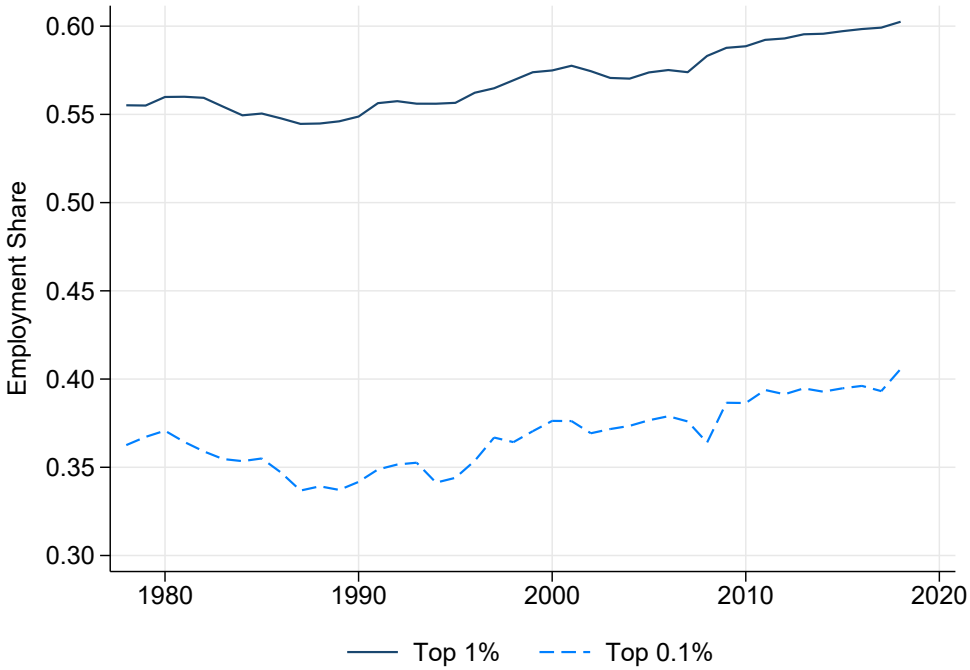


Figure IA12. Employment Concentration in Census Business Dynamic Statistics

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 data. See Internet Appendix IA2.3 for details about variable construction.

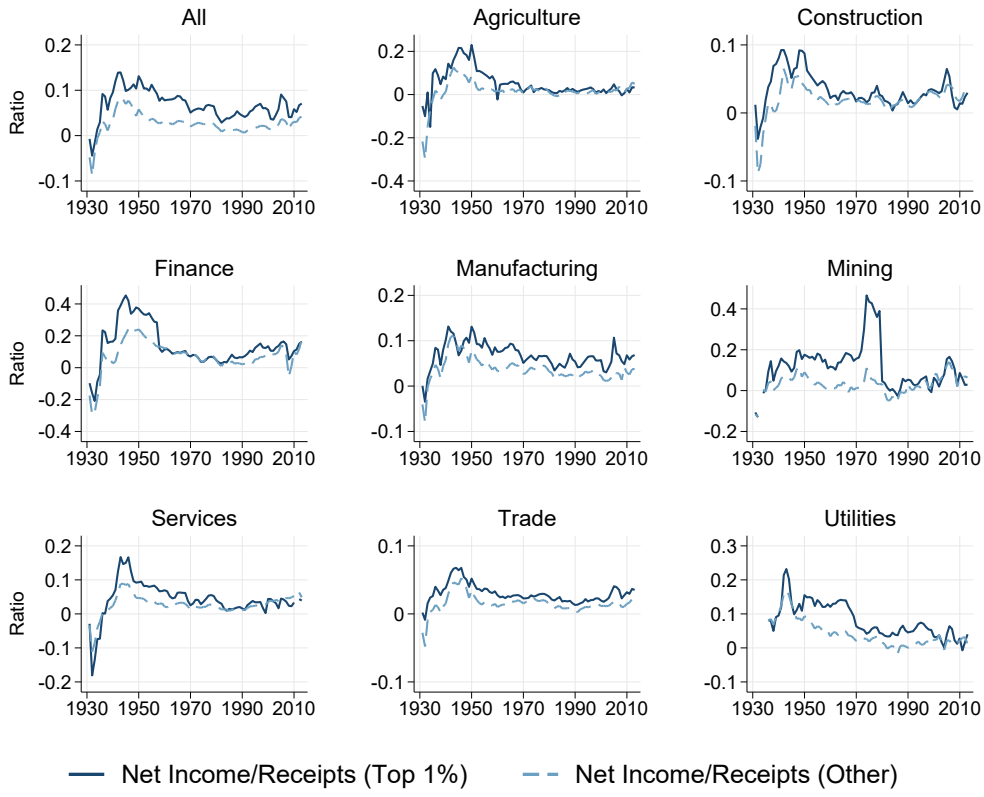


Figure IA13. Profitability

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 I 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.

Figure IA14. 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.

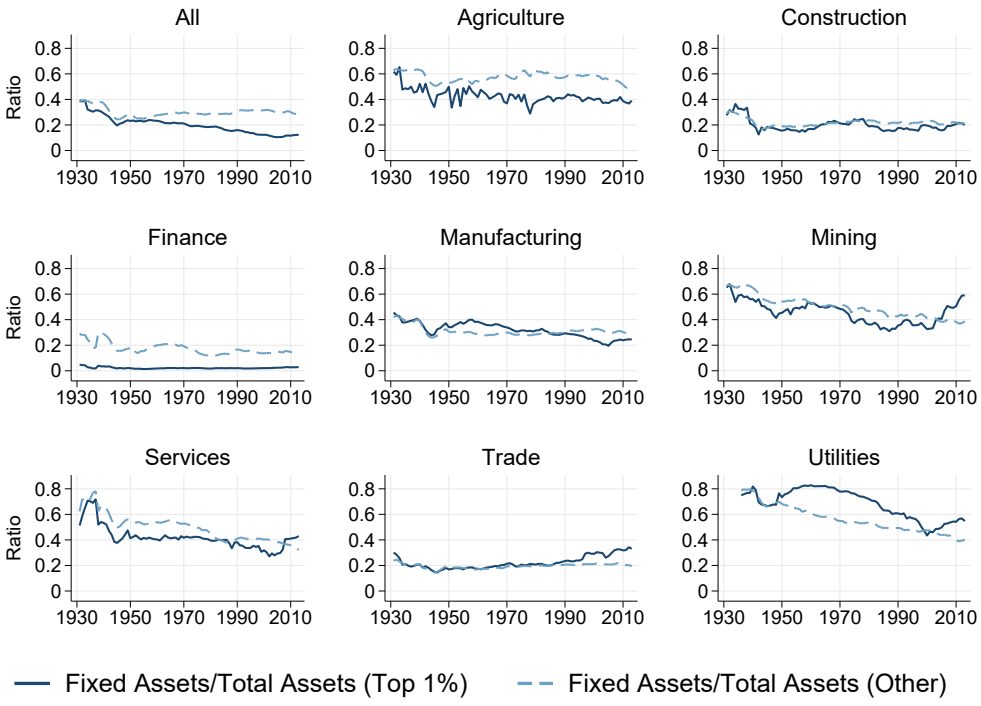


Figure IA15. Fixed Assets/Total Assets

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 I 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.

IA2 Data Construction

IA2.1 SOI Data

We digitize annual data from historical publications of the Internal Revenue Service (IRS). The IRS has a long tradition of collecting detailed statistics for individuals and businesses going back to the Revenue Act of 1916. The Statistics of Income report (SOI) was first published in 1918. In addition to data on receipts and net income, the SOI contains data on balance sheets as well (from balance sheet information submitted in tax returns). Using micro data from these submissions, the SOI provides tabulations of businesses by size of net income since 1918 (which ended in 1974), by size of assets since 1931, and by size of receipts since 1959. We use these tabulations by size to study concentration trends over the long run. As discussed in Section I, the tabulations by size are most comprehensive for corporations (both C-corporations and S-corporations). We provide additional checks for concentration estimates including noncorporations in Section II when we can obtain relevant data, and explain these data on noncorporations at the end of this section.

The SOI publications are accompanied by the Corporation Source Book, which is a series of initially unpublished volumes containing tabulations of corporations by asset size. The Corporation Source Book has more granular data and more detailed industry classifications compared to the main SOI reports. The Corporation Source Book is available through the IRS website (Internal Revenue Service, 2024*c*,*e*) from 1964 (digitally after 2000), and through the Electronic Records Division at the U.S. National Archives and Records Administration (National Archives and Records Administration, 2024) for certain years. We use the Corporation Source Book whenever available.

We provide links to all the sources for each year's data in the Excel spreadsheet "SOI Sources" in our replication package (Kwon, Ma, and Zimmermann, 2024).

Scope The scope of the SOI tabulations of corporations by size 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).[1] 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.[2] Starting in 1951, the IRS began to use a stratified probability sample to provide estimates for the whole population.[3] 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

[1] 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.

[2] See for example the 1965 SOI report Historical Summary (for the 50th year of SOI reports) on page 333: https://www.irs.gov/pub/irs-soi/65cocrar.pdf. "Small corporations whether liable for the tax or taxed at the shareholder level have always been included." "For 1916-1950 data were extracted from each return filed. Beginning with 1951, universe estimates were obtained from scientific samples."

[3] See for example the 1951 SOI report page 3: https://www.irs.gov/pub/irs-soi/51soireppt2ar.pdf.

reliable totals. The sample usually included the universe of businesses in the top bins. Therefore, the transition to sample data should not affect our concentration measures.

A small fraction of corporations do not submit information about their balance sheets. These cases are usually corporations without assets (e.g., liquidations, dissolutions, and 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 report of 1958, these filings were included in all tabulations by net income, but excluded from tables pertaining to balance sheet information.[4] Starting in the SOI report of 1959, the IRS included corporations with zero assets in the balance sheet tabulations and imputed data for corporations with missing balance sheets using information from the returns of corporations with both income statements and balance sheets in the same industry.[5] The left panel of Figure IA16 shows the share of corporation returns in each year with balance sheet information and the receipt share accounted for by these returns. For example, in 1950 about 10% of corporation returns representing 1.2% of receipts did not include balance sheet data. Both the share of returns without balance sheet data and their receipt share declined over time. For the tabulations by assets before 1959 (which did not include corporations without balance sheet data), we can provide robustness checks for our top share estimates by either assuming that the corporations 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 IA16 compares our baseline concentration estimates to concentration estimates with imputed assets for corporations with missing balance sheet information. The two series are similar. We also find the same degree of consistency at the industry level (results not shown).

From 1959 to 1964, the tabulations of corporations by receipts have a category "No business receipts reported," which contains a small fraction of corporations. We currently exclude them in our analyses, to be consistent with our treatment of corporations with missing balance sheet data discussed above. This category does report total receipts, which appear to be very small. The "No business receipts reported" category is not present in the by receipt size tabulations starting in 1965.

Variable definition For definitions of "total assets," "total receipts," "business receipts," "net income," and other terms, see for example the "Explanations of Terms" section of the 2013 SOI report: https://www.irs.gov/pub/irs-soi/13coccr.pdf. The tabulations by receipt size use the size of business receipts for the aggregate and main sectors, and the size of total receipts for "Finance, Insurance, and Real Estate" during the SIC years and "Finance and Insurance" and "Management of Companies (Holding Companies)" during the NAICS years.

Industry classification The SOI assigns a single industry code to each business based on the industry that represents the largest percentage of its receipts.[6] 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

[4] See for example the 1958 SOI report page 9: https://www.irs.gov/pub/irs-soi/58cocrar.pdf.

[5] See for example the 1959 SOI report pages 16 to 17: https://www.irs.gov/pub/irs-soi/59cocrar.pdf.

[6] See for example the 1931 SOI report page 4: https://www.irs.gov/pub/irs-soi/31soirepar.pdf; the 1977 SOI report page 15: https://www.irs.gov/pub/irs-soi/77cocrar.pdf.

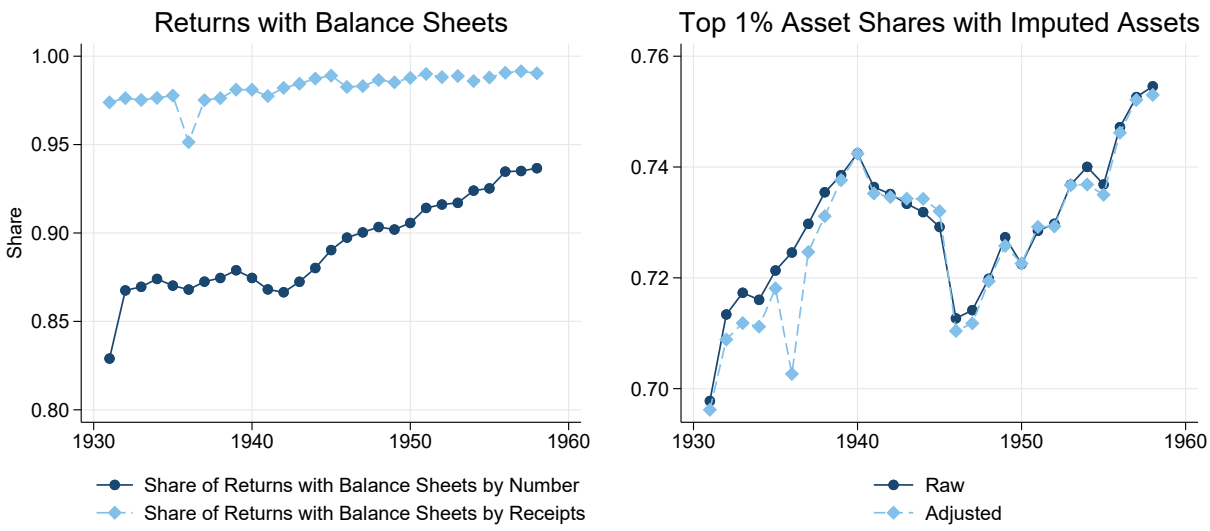


Figure IA16. Corporation Returns without Balance Sheets

Notes: The left panel shows the share of corporation returns with balance sheet information in all corporation returns, by number (dark blue circles) and by receipts (light blue diamonds). The right panel shows top asset shares if we impute assets for corporations 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.

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 industry groupings have remained relatively stable over 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).

We use two main industry harmonization schemes. The first scheme applies to tabulations by assets, where we have more granular industry breakdown (e.g., from the Corporation Source Book). Here we can handle the transition between SIC and NAICS in a more detailed way. Table IA1, 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. 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: Other," 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, NIACS 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).

The second and coarser scheme applies to tabulations by receipts, where the industry breakdown is less granular around the SIC to NAICS transition. For these tabulations, the mapping to the main sectors is as follows. Agriculture: SIC 01-09; NAICS 11. Mining: SIC 10-14; NAICS 21. Construction: SIC 15-17; NAICS 23. Manufacturing: SIC 20-39; NAICS 31-33. Utilities: SIC 40-49; NAICS 22, 48-49, and 51. Trade: SIC 50-59; NAICS 42-45. Finance: SIC 60-67; NAICS 52, 53, and 55. Services: SIC 70-89; NAICS 54, 56, 61, 62, 71, 72, and 81.

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 years, we can manually back out the missing values using the adding up constraints from the hierarchical industry and bin structure (similar in spirit to Eckert et al., 2020). In later years, further precautions have been introduced by the IRS to preserve confidentiality by deleting information from additional bins when necessary. In these cases, we combine the deleted bins (and all bins in between) into one large bin, and back out the financial data using the difference between the total and all other bins. While this approach generally works well, in a handful of cases the number of size bins is reduced too much to obtain consistent and robust top shares, and we linearly interpolate top share estimates for these years.

Consolidation Corporations can file consolidated returns if at least 80% of the equity of an affiliate is owned within the group. The SOI reports consolidated affiliates as one entity. Corporations that choose to file consolidated returns in one year are generally also required to file consolidated returns in 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.[7] In recent years at least, eligible firms generally elect to consolidate (Mills, Newberry, and Trautman, 2002), given more favorable treatment when consolidated (e.g., 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, and we provide a summary here. First, the 80% ownership requirement applicable today dates back to 1954. Prior to 1954, the ownership threshold was 95%.[8] Second, consolidated returns were often taxed at higher rates before the 1960s. In 1932 and 1933, consolidated returns were subject to an additional tax of 0.75%.[9] In 1934 and 1935, the additional tax increased to 1%.[10] No additional tax was imposed between 1936 and 1941, but the consolidation privilege was significantly limited (see below). Between 1942 and 1963, corporations

[7] See for example the 2013 SOI report page 269: https://www.irs.gov/pub/irs-soi/13coccr.pdf.
[8] See for example the 1954 SOI report page 14: https://www.irs.gov/pub/irs-soi/59cocrar.pdf.
[9] See for example the 1932 SOI report page 1: https://www.irs.gov/pub/irs-soi/32soirepar.pdf.
[10] See for example the 1934 SOI report page 30: https://www.irs.gov/pub/irs-soi/34soireppt2ar.pdf.

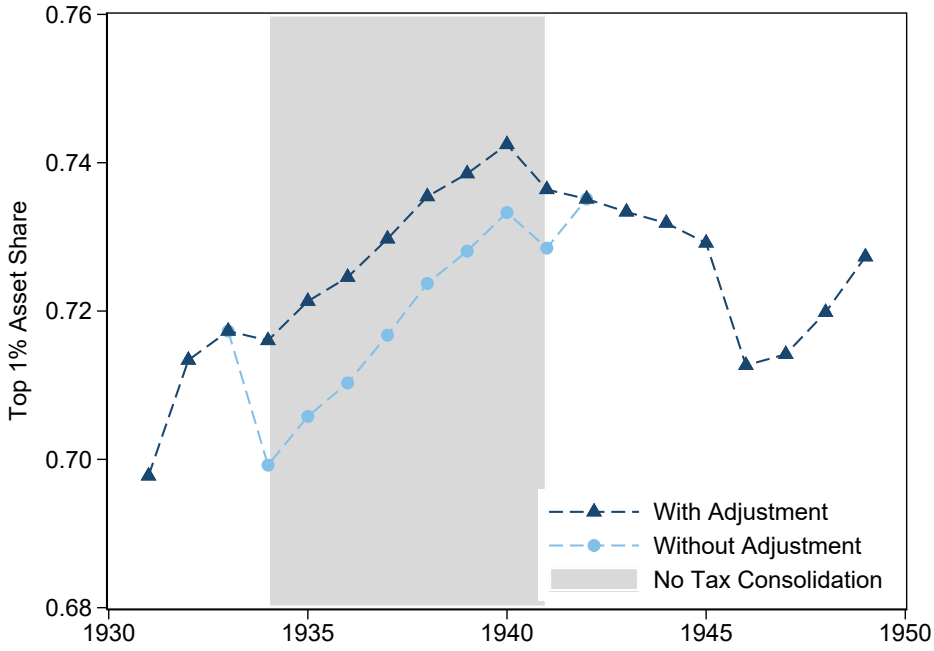


Figure IA17. Consolidation Adjustment

Notes: This figure shows the top 1% asset share between 1931 and 1949 with and without adjustment for changes in consolidation.

filing consolidated returns were subject to a surtax on the group of 2%.[11] This 2% surtax was repealed by the Revenue Act of 1964, and no such surtax has been applied since then. Third, consolidation was mandatory between 1918 and 1921, and voluntary after 1922.[12] Then between 1934 and 1941, all corporations (except for railway companies that were affiliated with each other) were not allowed to file consolidated returns.[13] This policy change only induced a modest decline in the top 1% asset share in the aggregate (see Figure IA17), though 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 IA17. The dashed line with circles shows the top 1% asset shares without adjustment, and the dashed line with triangles shows the adjusted series.

[11] See for example the 1943 SOI report page 4: https://www.irs.gov/pub/irs-soi/42soireppt2ar.pdf.
[12] See for example the 1918 SOI report page 12: https://www.irs.gov/pub/irs-soi/18soirepar.pdf.
[13] See for example the 1934 SOI report page 19: https://www.irs.gov/pub/irs-soi/34soireppt2ar.pdf.

One possible concern is that changes in the prevalence of consolidation may affect the concentration trends we observe. We perform several checks. First, we compare our results with information from companies' consolidated financial statements. 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 after the adjustment. Their results are similar to our estimates of the top 0.1% share among nonfinancial corporations for those years. In addition, for manufacturing where rising concentration in the early decades is most prominent, Figure IA9 shows that our top share estimates using SOI data are similar to estimates using assets from top firms' consolidated financial statements. For more recent decades, the comparison with Compustat data in Figure IA10 also shows that top businesses in SOI data are similar to those in financial statements.[14]

Second, we digitize data on consolidated returns in SOI reports. Figure IA18 shows the share of consolidated returns in corporation returns by number (dark blue circles), and the share of assets from consolidated returns (light blue diamonds). We observe a decrease in the prevalence of consolidated returns between the early 1930s and the 1940s, and an increase from the mid-1960s to the 1980s, roughly returning to the prevalence of consolidated returns in the early 1930s. Meanwhile, the 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 assets, while the top 1% shares continued to rise. Figure IA19 also shows that we generally observe rising top shares within each period of consolidation rules, using the manufacturing and aggregate series as examples (here we present the final top 1% asset share estimates, and the only modification to the raw estimates from SOI data is the adjustment for the 1934 to 1941 period explained above).

Third, the consolidation rules apply to all sectors and the consolidation trends are similar across sectors, but the concentration trends display differences in the timing of rising concentration.

Partnerships and sole proprietorships We collect three types of data for noncorporations (partnerships and sole proprietorships) from historical SOI publications. First, we collect information on the number of noncorporations, as well as the share of corporations in corporate and noncorporate receipts. We use the dataset compiled by Lamoreaux (2006) up to 1997, and we extend the series afterwards (Internal Revenue Service, 2024*a*).

Second, we are able to obtain tabulations of noncorporations by size bins of receipts for the main sectors, from 1959 to 1980 (Internal Revenue Service, 2024*d*) and 1998 to 2003 (Internal Revenue Service, 2024*a*). Table IA2 shows the detailed list of these sources. The partnership tabulations separately report the number of returns with "Receipts not reported" from 1959 to 1962 and 1965 to 1968, and "No receipts reported" from 1973 to 1979. Here we treat them as zero receipts to maintain consistency across the years, since the tabulations in 1972 do not have such a category but note that the smallest bin "Includes partnerships without business receipts." We do the same for the "Receipts not reported"/"No receipts reported" category in the proprietorship tabulations.

[14]For financial accounting, the ownership threshold for consolidation is 50%, which is lower than the 80% threshold for tax filing. Mills, Newberry, and Trautman (2002) suggest that large businesses do not have many subsidiaries with ownership between 50% and 80%.

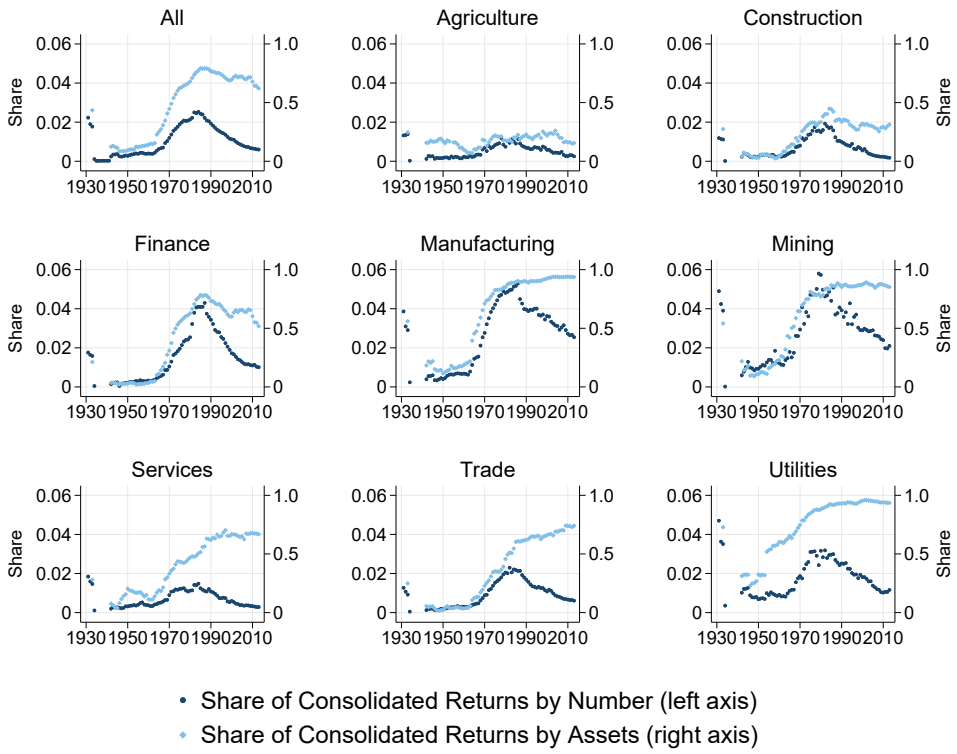


Figure IA18. Prevalence of Consolidation

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

Third, in the aggregate, we are able to obtain tabulations of partnerships by size bins of assets from 1965 to 1982 (Internal Revenue Service, 2024*d*) and 2002 onward (Internal Revenue Service, 2024*b*). Table IA3 shows the detailed list of these sources. We use these data in Figure 8 to account for large noncorporations in the top N asset share. Most large noncorporations are partnerships, so the asset size distribution of partnerships is helpful, especially for recent years as large noncorporations have become more common. The early tabulations (between 1965 and 1982) have a category "Partnerships without balance sheets," which we currently exclude. This does not affect our results because we only need information about the largest partnerships in Figure 8; moreover, large partnerships were uncommon in this early period, in which case the top N asset share relies little on partnership tabulations.

Data on partnerships come from Form 1065, which partnerships need to file with the IRS for informational purposes. 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 receipts of proprietors, rather than those of individual proprietorships. However, the ratio between the number of companies and the number of owners is close to 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

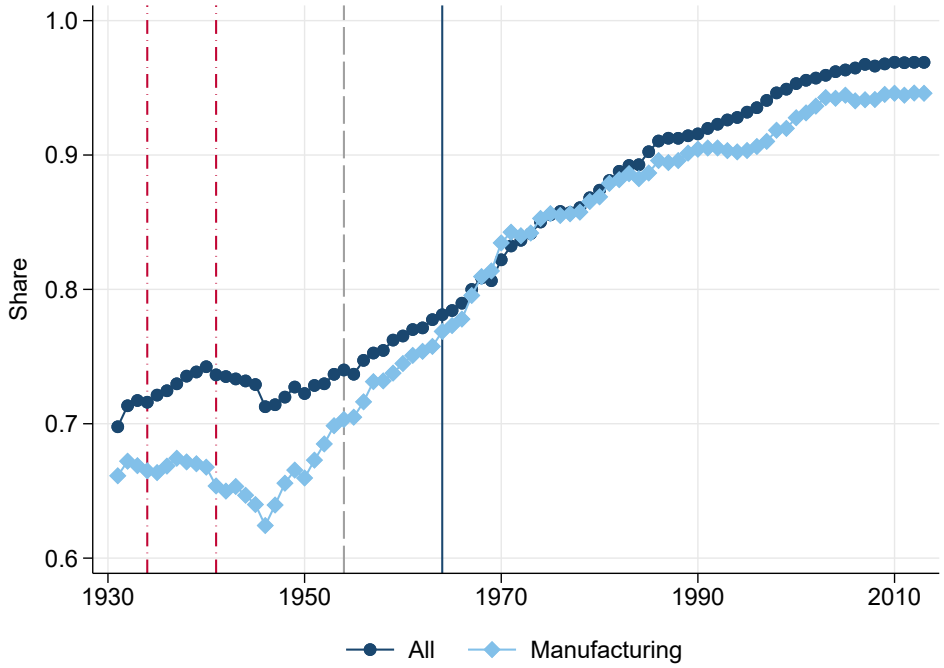


Figure IA19. Top 1% Asset Shares under Different Consolidation Rules

Notes: This figure shows the top 1% asset share for the aggregate (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 ownership threshold for consolidation changed from 95% to 80%. The solid blue line marks 1964, where the surtax on consolidated returns ended.

some years, the reporting unit was the number of C Schedules filed with the return, and in some years the SOI only counted those businesses that operated in different industries as separate businesses, or restricted the overall number of businesses per owner (Lamoreaux, 2006).

Table IA1 – Detailed 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)	521, 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), Petroleum (29)	324, 325
Manufacturing: Electrical	Electronic (36)	334, 335
Manufacturing: Food	Food and Kindred Products (20), Tobacco Products (21)	311, 312
Manufacturing: Machinery	Industrial and Commercial Machinery (35), Measuring, Analyzing, and Controlling Instruments (38)	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	Wholesae 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. This detailed industry harmonization scheme is used for tabulations of corporations by asset size. SOI industries are classified by economic activity using SIC codes until 1997 and NAICS 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 industry codes.

Table IA2 – Sources for Noncorporation Tabulations by Receipts

Year	Source	Business Type
1959	U.S. Business Tax Returns, 1959-1960 (pages 18ff. & 46ff.)	Sole Proprietorships, Partnerships
1960	U.S. Business Tax Returns, 1960-1961 (pages 24ff. & 56ff.)	Sole Proprietorships, Partnerships
1961	U.S. Business Tax Returns, 1961-1962 (pages 24ff. & 54ff.)	Sole Proprietorships, Partnerships
1962	U.S. Business Tax Returns, 1962 (pages 34ff. & 120ff.)	Sole Proprietorships, Partnerships
1964	U.S. Business Tax Returns, 1964 (pages 87ff.)	Partnerships
1965	U.S. Business Income Tax Returns, 1965 (pages 39ff. & 129ff.)	Sole Proprietorships, Partnerships
1966	U.S. Business Income Tax Returns, 1966 (pages 37ff. & 130ff.)	Sole Proprietorships, Partnerships
1967	U.S. Business Income Tax Returns, 1967 (pages 41ff. & 165ff.)	Sole Proprietorships, Partnerships
1968	U.S. Business Income Tax Returns, 1968 (pages 36ff. & 143ff.)	Sole Proprietorships, Partnerships
1969	U.S. Business Income Tax Returns, 1969 (pages 34ff. & 115ff.)	Sole Proprietorships, Partnerships
1970	U.S. Business Income Tax Returns, 1970 (pages 38ff. & 123ff.)	Sole Proprietorships, Partnerships
1971	U.S. Business Income Tax Returns, 1971 (pages 35ff. & 123ff.)	Sole Proprietorships, Partnerships
1972	U.S. Business Income Tax Returns, 1972 (pages 18ff. & 102ff.)	Sole Proprietorships, Partnerships
1973	U.S. Business Income Tax Returns, 1973 (pages 23ff. & 144ff.)	Sole Proprietorships, Partnerships
1974	U.S. Business Income Tax Returns, 1974 (pages 30ff. & 145ff.)	Sole Proprietorships, Partnerships
1975	U.S. Business Income Tax Returns, 1975 (pages 22ff. & 220ff.)	Sole Proprietorships, Partnerships
1976	U.S. Business Income Tax Returns, 1976 (pages 34ff. & 259ff.)	Sole Proprietorships, Partnerships
1977	Sole Proprietorship Returns, 1977 (pages 42ff.)	Sole Proprietorships
1977	Partnership Returns, 1977 (pages 29ff.)	Partnerships
1978	Sole Proprietorship Returns, 1978 (pages 27ff.)	Sole Proprietorships
1978	Partnership Returns, 1978 (pages 27ff.)	Partnerships
1979	Sole Proprietorship Returns, 1979-1980 (pages 34ff.)	Sole Proprietorships
1979	Partnership Returns, 1979 (pages 29ff.)	Partnerships
1980	Sole Proprietorship Returns, 1979-1980 (pages 165ff.)	Sole Proprietorships
1980	Partnership Returns, 1980 (pages 33ff.)	Partnerships
1998	IRS Website, Integrated Business Data	Sole Proprietorships, Partnerships
1999	IRS Website, Integrated Business Data	Sole Proprietorships, Partnerships
2000	IRS Website, Integrated Business Data	Sole Proprietorships, Partnerships
2001	IRS Website, Integrated Business Data	Sole Proprietorships, Partnerships
2002	IRS Website, Integrated Business Data	Sole Proprietorships, Partnerships
2003	IRS Website, Integrated Business Data	Sole Proprietorships, Partnerships

Notes: This table shows the sources for the tabulations of size bins by receipts for noncorporations. We exclude the main sector "Finance" for 1998 to 2003 due to an inconsistency in the tabulations for partnerships confirmed by IRS staff: the column on "gross receipts" has data errors, and the values presented there do not represent gross receipts.

Table IA3 – Sources for Partnership Tabulations by Assets

Year	Source	Business Type
1965	U.S. Business Income Tax Returns, 1965 (pages 197ff.)	Partnerships
1966	U.S. Business Income Tax Returns, 1966 (pages 161ff.)	Partnerships
1967	U.S. Business Income Tax Returns, 1967 (pages 207ff.)	Partnerships
1968	U.S. Business Income Tax Returns, 1968 (pages 186ff.)	Partnerships
1969	U.S. Business Income Tax Returns, 1969 (pages 157ff.)	Partnerships
1970	U.S. Business Income Tax Returns, 1970 (pages 151ff.)	Partnerships
1971	U.S. Business Income Tax Returns, 1971 (pages 153ff.)	Partnerships
1972	U.S. Business Income Tax Returns, 1972 (pages 96ff.)	Partnerships
1973	U.S. Business Income Tax Returns, 1973 (pages 208ff.)	Partnerships
1974	U.S. Business Income Tax Returns, 1974 (pages 199ff.)	Partnerships
1975	U.S. Business Income Tax Returns, 1975 (pages 310ff.)	Partnerships
1976	U.S. Business Income Tax Returns, 1976 (pages 395ff.)	Partnerships
1977	Partnership Returns, 1977 (pages 113ff.)	Partnerships
1978	Partnership Returns, 1978 (pages 75ff.)	Partnerships
1979	Partnership Returns, 1979 (pages 45ff.)	Partnerships
1980	Partnership Returns, 1980 (pages 49ff.)	Partnerships
1981	Partnership Returns, 1978-1982 (pages 153ff.)	Partnerships
1982	Partnership Returns, 1978-1982 (pages 197ff.)	Partnerships
2002	IRS Website, Partnership Statistics	Partnerships
2003	IRS Website, Partnership Statistics	Partnerships
2004	IRS Website, Partnership Statistics	Partnerships
2005	IRS Website, Partnership Statistics	Partnerships
2006	IRS Website, Partnership Statistics	Partnerships
2007	IRS Website, Partnership Statistics	Partnerships
2008	IRS Website, Partnership Statistics	Partnerships
2009	IRS Website, Partnership Statistics	Partnerships
2010	IRS Website, Partnership Statistics	Partnerships
2011	IRS Website, Partnership Statistics	Partnerships
2012	IRS Website, Partnership Statistics	Partnerships
2013	IRS Website, Partnership Statistics	Partnerships

Notes: This table shows the sources for the tabulations of size bins by assets for partnerships.

IA2.2 Additional Historical Data on Manufacturing Top 100 Asset Share

Several sources compiled the asset size of the largest firms for the early 1900s from financial statements, since the reporting of assets was more common than the reporting of sales at that time (Navin, 1970). We use these sources to obtain the share of the largest 100 manufacturers by assets in total manufacturing assets in Figure IA9.

SOI data The benchmark series is the top 100 asset share in manufacturing estimated using our SOI data. The top bins by assets are very granular before the 1980s (containing just around 100 corporations in the manufacturing tabulations in these early years), so we can estimate the assets of the top 100 fairly well in the early decades; we can also reasonably assume that the largest 100 manufacturers are predominantly corporations.

For total business assets in the denominator of the top shares, we start with the assets of corporations and estimate the assets of noncorporations. In Figure 8, we approximate the ratio of corporate to noncorporate assets using the ratio of corporate to noncorporate receipts from SOI data (starting from 1945). To cover the early decades of the 20th century, here we use the ratio of corporate to noncorporate value added from historical manufacturing census data (linearly interpolated between census years) in Atack and Bateman (2024). Panel B of Figure IA6 shows that corporations' share in corporate and noncorporate value added based on manufacturing census data is similar to corporations' share in corporate and noncorporate receipts based on SOI data.[15] This procedure to estimate total manufacturing assets is used in all series except for the Federal Trade Commission (1969) series, which has already estimated noncorporate assets for the denominator of their top shares using a similar approach.

We also provide the estimated share of the top 100 manufacturers by net income. Here the numerator is the net income of the 100 largest manufacturers by net income; the denominator is the net income of manufacturing corporations (restricted to those with positive net income) and the corresponding estimated net income of noncorporations (using the corporate share in manufacturing value added as explained above).

Federal Trade Commission publications Federal Trade Commission (1969) provides estimates for the asset share of the largest 100 manufacturers from the Federal Trade Commission's historical congressional hearings. They obtain the numerator from *Moody's Industrial Manual*. Their denominator sums up corporate assets in SOI data with noncorporate assets estimated using SOI and census data (see Appendix B of Federal Trade Commission (1969) for details). We transcribe their estimates in Part 8A Table 3-3 "Share of manufacturing assets held by the 200 largest corporations, 1925-41; 1947-48."

List of largest industrials by Collins and Preston (1961) Collins and Preston (1961) provide a list of the largest 100 industrial firms (predominantly manufacturing with some mining) for 1909, 1919, 1929, 1935, 1948, and 1958. We transcribe the data from their table "The 100 Largest Industrial Firms, Selected Years, 1909-1958." We sum the assets of these companies in the numerator. We use SOI data to estimate the denominator following the procedures discussed above: we start with corporate assets in manufacturing and mining (manufacturing accounts for about 80% to 90% of the sum), and then estimate noncorporate assets using the corporate value added share in manufacturing (we have to assume that the corporate share is similar in mining).[16]

[15] We extend the corporate share in manufacturing value added in Atack and Bateman (2024) with data from 1997 to 2012, using U.S. Census Bureau (1997) (Table 3-3 "Industry Statistics for Subsectors by Type of Company Ownership and Legal-Form of Organization: 1997"), U.S. Census Bureau (2002) (Table 11, "Industry Statistics for Subsectors by Legal Form of Organization: 2002"), U.S. Census Bureau (2007) (Table EC0731SG6 "Manufacturing: Summary Series: General Summary: Industry Statistics for Subsectors by Legal Form of Organization: 2007"), and U.S. Census Bureau (2012) (Table EC1231SG5 "Manufacturing: Summary Series: General Summary: Industry Statistics for Subsectors by Legal Form of Organization: 2012").

[16] The SOI has published detailed sector-level balance sheets since 1926. In addition, the SOI published total corporate

List of largest manufacturers by Chandler (1994) Chandler (1994) provides a list of the largest 200 manufacturers for 1917, 1930, and 1948. We transcribe the data from Appendix Table A.1, A.2, and A.3. We sum the assets of the top 100 companies for the numerator. We use SOI data to estimate total manufacturing assets as discussed above.[17]

IA2.3 Variable Construction

In the following we explain our variable construction process in detail.

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/Assets of U.S. corporations in a given year. Used in Table 3, Figure 1, Figure 3, Panel A of Figure 4, Figure 5, Figure 8, Figure IA1, Figure IA3, Figure IA4, and Figure IA11.
   - Top x% receipt share = Receipts of top x% corporations by receipts in a given year/Receipts of U.S. corporations in a given year. Used in Figure 1, Figure 3, Panel B of Figure 4, Figure 6, Figure 7, and Figure IA8.
   - Top x% net income share = Net income of top x% corporations by net income among corporations with positive net income in a given year/Net income of U.S. corporations with positive net income in a given year. Used in Figure 1 and Figure 3.
   - Top x% equity capital share = Equity capital of top x% corporations by equity capital in a given year/Equity capital of U.S. corporations in a given year. Used in Figure 1.

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

equity by sector for 1909 to 1913 (Annual Report of the Commissioner of Internal Revenue, 1910–1914) and for 1921, 1922, and 1924 (SOI 1920, 1922, and 1925). To obtain total assets for industrials prior to 1926, we collect total corporate equity for industrials (linearly interpolating the gaps), and scale it with the 1926 capital ratio (0.71). The Collins and Preston (1961) list has some nonindustrial companies. These firms account for a small fraction (2% to 5%) of the top 100 assets in their data.

[17] The denominator in this case is identical to that of Collins and Preston (1961), except that it uses manufacturing assets rather than the assets of industrials (manufacturing and mining). No breakdown of industrials is available prior to 1921. We estimate total manufacturing assets prior to 1921 using the 1921 ratio of manufacturing to industrial fair value capital stock (0.82).

Procedure We examine several methods to estimate top shares. Results for these methods are over 0.99 correlated as explained in Section I 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 corporations 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 corporations in these bins approximates $x\%$ (as long as the first bin contains less than $x\%$ of corporations). For example, if the total number of corporations is N and the number of corporations 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 corporations 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 corporations in the ith bin). In other words, we take all corporations in the top k bins and fill in the residual from the $k+1$th bin.

2. Relative concentration among top corporations

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 and Figure IA2.
- 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.
- Top $x\%$ equity capital share among top $y\%$ = Equity capital of top $x\%$ corporations by equity capital in a given year/Equity capital of top $y\%$ corporations by equity capital in a given year. Used in Figure 2.

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

3. Top $x\%$ receipt share among corporations plus noncorporations

Definition Top $x\%$ receipt share = Receipts of top $x\%$ businesses by receipts (among corporations and noncorporations) in a given year/Receipts of corporations and noncorporations in a given year. Used in Figure 6, Figure 7, and Figure IA8.

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 distributions using the gpinter routine of Blanchet, Fournier, and Piketty (2022) to obtain estimates for top receipt shares among all businesses.

4. Top N receipt share among corporations plus noncorporations

Definition Top N receipt share = Receipts of top N businesses by receipts (among corporations and noncorporations) in a given year/Receipts of corporations and noncorporations in a given year. Used in Figure 7, Figure IA7, and Figure IA8.

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 distributions using the gpinter routine of Blanchet, Fournier, and Piketty (2022) to obtain estimates for top receipt shares among all businesses.

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 assets of corporations plus noncorporations in a given year. Used in Figure 8 and Figure IA9.
- Top N receipt share = Receipts of top N corporations by receipts in a given year/Receipts of corporations plus noncorporations in a given year. Used in Panel B of Figure 7.
- 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 net income of corporations plus noncorporations (with positive net income) in a given year. Used in Figure IA9.

Procedure For total business assets in the denominator of the top N asset share, we need to estimate noncorporate assets. For the aggregate and main sectors after 1945 (Figure 8), we scale corporate assets with the ratio of corporate and noncorporate receipts using SOI data:

$$\widehat{\text{Noncorp assets}}_t = \text{Corp assets}_t \times \frac{\text{Noncorp receipts}_t}{\text{Corp receipts}_t}.$$

This approach assumes that the ratio of noncorporate to corporate assets and the ratio of noncorporate to corporate receipts are comparable. We check this assumption in Panel A of Figure IA6 using information about corporate and noncorporate assets from the Flow of Funds, available for nonfinancial industries in the aggregate. We verify that corporations' share in corporate and noncorporate receipts is similar to corporations' share in corporate and noncorporate assets. For manufacturing in the early 20th century (Figure IA9), we scale corporate assets with the ratio of corporate and noncorporate value added using manufacturing census data:

$$\widehat{\text{Noncorp assets}}_t = \text{Corp assets}_t \times \frac{\text{Noncorp value added}_t}{\text{Corp value added}_t}.$$

Panel B of Figure IA6 shows that corporations' share in corporate and noncorporate value added in manufacturing is similar to corporations' share in corporate and noncorporate receipts (based on SOI data after 1945), and earlier we checked that corporations' share in corporate and noncorporate assets can be approximated by their share in receipts.

For the top N receipt share, we multiply the top N corporate receipt share with the ratio of corporate to noncorporate receipts from SOI data. For the top N net income share, we multiply the top N corporate net income share with the ratio of corporate to noncorporate value added for manufacturing in the early 20th century (Figure IA9).

6. Top N corporations' and partnerships' asset share among corporations plus noncorporations

Definition Assets of top N corporations and partnerships by assets in a given year/Estimated assets of corporations and noncorporations in a given year. Used in Figure 8.

Procedure We estimate this ratio in years where we have tabulations of partnerships by size bins based on assets. We first estimate separate distributions for corporations and partnerships, and then merge these two distributions using the gpinter routine of Blanchet, Fournier, and Piketty (2022) to obtain estimates for the assets of the largest N corporations and partnerships. We then normalize this with the estimated assets of corporations and noncorporations as described above (assuming that the ratio of noncorporate to corporate assets is comparable to the ratio of noncorporate to corporate receipts).

7. Top N corporations' share among corporations

Definition Top N receipt share among corporations = Receipts of top N corporations by receipts in a given year/Receipts of U.S. corporations in a given year. Used in Panel B of Figure 7.

Procedure We use the generalized Pareto interpolation method described above.

8. Receipt share of top 20 corporations by assets to compare with census concentration ratios

Definition Top 20 receipt share in SOI data = Receipts of top 20 corporations by assets in a given year/Receipts of corporations in a given year. Used in Panel B of Figure 9 (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 α fraction of the $k+1$th bin, such as the number of corporations in the top k bins plus α fraction of the number of corporations in the $k+1$th bin is 20. We sum the receipts of corporations in the top k bins and α fraction of the receipts of the $k+1$th bin for the numerator. (Here we cannot use the Pareto interpolation because we need the receipts of the top businesses by *assets*, given that only size bins by assets are available for these granular industries.)

9. Top 1% asset 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 foreign affiliates)/(Assets of U.S. corporations + assets of the foreign affiliates of U.S. multinationals). Used in Figure IA11.

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 foreign 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 corporations in the numerator (i.e., add all international assets to the numerator); 2) assume the top corporations' 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 1% asset share among U.S. corporations).

10. Profitability (net income before tax/sales)

Definition

- Profitability among top 1% corporations by assets = Net income before tax among top 1% corporations by assets in a given year/Total receipts among top 1% corporations by assets in a given year. Used in Figure IA13.
- Profitability among bottom 99% corporations by assets = Net income before tax among bottom 99% corporations by assets in a given year/Total receipts among bottom 99% corporations by assets in a given year. Used in Figure IA13.
- Profitability among all corporations (SOI) = Net income before tax among all U.S. corporations in a given year (SOI)/Total receipts among all U.S. corporations in a given year (SOI). Used in Figure IA14.

- Profitability among all corporations (BEA) = Corporate profits before tax in a given year (BEA)/Total receipts among all U.S. corporations in a given year (SOI). Used in Figure IA14.

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 the top 1% corporations (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 corporations in these bins approximates 1% (as long as the first bin contains less than 1% of corporations). For instance, if the total number of corporations is N and the number of corporations 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 corporations 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 corporations in the ith bin). We do so to estimate the net income and receipts of the top 1% corporations, and the remaining amount of net income and receipts belong to the bottom 99% corporations.

11. Fixed assets/total assets

Definition

- Fixed assets over total assets among top 1% corporations by assets = Fixed assets among top 1% corporations by assets in a given year/Total assets among top 1% corporations by assets in a given year. Used in Figure IA15.
- Fixed assets over total assets among bottom 99% corporations by assets = Fixed assets among bottom 99% corporations by assets in a given year/Total assets among bottom 99% corporations by assets in a given year. Used in Figure IA15.

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 corporations in these bins approximates 1% (as long as the first bin contains less than 1% of corporations). For instance, if the total number of corporations is N and the number of corporations 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 corporations 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 corporations in the ith bin). We do so to estimate the fixed assets and total assets of the top 1% corporations, and the remaining amount of fixed assets and total assets belong to the bottom 99% corporations.

Other Data

1. Concentration ratios 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 9.
   - CR20 among manufacturing and non-manufacturing industries in 2012. Used in Panel B of Figure 9.

   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://www2.census.gov/programs-surveys/economic-census/data.

   As explained in census documentations, the concentration ratios for each industry are based on the sales of each firm's establishments in that industry.[18] 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 I, the SOI data always treat a business as one observation, which belongs to the industry that represents the largest percentage of its total receipts.

2. Top share using Compustat data

   Definition

   - Compustat top 500 asset share = Assets of top 500 Compustat firms by assets in a given year/Total corporate assets in a given year. Used in Figure IA10.
   - Compustat top 500 sales share = Sales of top 500 Compustat firms by sales in a given year/Total corporate receipts in a given year. Used in Figure IA10.

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

3. Employment concentration in census BDS data

[18] "Each establishment is assigned a separate industry classification based on its primary activity and not that of its parent company." "A firm is a collection of one or more establishments under common ownership or operational control." "For each multi-establishment firm, establishments in the same industry within a geographic area will be counted as one firm."

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 IA12.

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.

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 $0 = t_0 < t_1 < t_2 < ....t_K$, 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 (μ_k, σ_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 Φ is the standardized normal CDF. The above equation translates into two linear equations in σ_k and μ_k, which can be solved exactly:

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

For the largest size bin (businesses with size above t_K), we fit the lognormal parameters (μ_K, σ_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 μ_K, we use the equation

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

to pin down the implied σ_K, and then confirm whether:

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

Lastly, for the smallest bin (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 μ_0. For this, we use the fact that $\int_a^b x dF(x)$ and $\int_a^b dF(x)$ for $F \sim$ Log Normal(μ, σ) 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:

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

Our fitted top share is given by $\frac{\hat{E}[x \geq t^*]}{\mu_0}$, where μ_0 is the overall mean.

References

Atack, Jeremy, and Fred Bateman. 2024. "Corporate Control of Manufacturing — Percentage of Workers and Value Added in Establishments Owned by Corporations: 1899-1992." , https://hsus.cambridge.org/HSUSWeb/toc/showTable.do?id=Dd875-918 (Accessed: April 30, 2024).

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty. 2022. "Generalized Pareto Curves: Theory and Applications." *Review of Income and Wealth*, 68(1): 263–288.

Chandler, Alfred D. 1994. *Scale and Scope: The Dynamics of Industrial Capitalism.* Harvard University Press.

Collins, Norman R., and Lee E. Preston. 1961. "The Size Structure of the Largest Industrial Firms, 1909-1958." *American Economic Review*, 51(5): 986–1011.

Eckert, Fabian, Teresa C. Fort, Peter K. Schott, and Natalie J. Yang. 2020. "Imputing Missing Values in the US Census Bureau's County Business Patterns." Working Paper.

Federal Trade Commission. 1969. "Economic Report on Corporate Mergers." *US Senate, Committee on the Judiciary, Subcommittee on Antitrust and Monopoly, Economic Concentration, Hearings Pursuant to S. Res*, 40.

Internal Revenue Service. 2024*a*. "SOI Tax Stats - Integrated Business Data." , https://www.irs.gov/statistics/soi-tax-stats-integrated-business-data (Accessed: September 26, 2020).

Internal Revenue Service. 2024*b*. "SOI Tax Stats - Partnership Data by Size of Total Assets." , https://www.irs.gov/statistics/soi-tax-stats-partnership-data-by-size-of-total-assets (Accessed: September 26, 2020).

Internal Revenue Service. 2024*c*. "SOI Tax Stats Archive - 1964 to 1999 Corporation Source Book." , https://www.irs.gov/statistics/soi-tax-stats-archive-1964-to-1999-corporation-source-book (Accessed: September 26, 2020).

Internal Revenue Service. 2024*d*. "SOI Tax Stats Archive, 1953 to 1984 - Business Income Tax Return Reports: Partnership, Sole Proprietorship, and Farmers' Cooperative Return Reports." , https://www.irs.gov/statistics/soi-tax-stats-archive-1953-to-1984-business-income-tax-return-reports-partnership-sole-proprietorship-and-farmers-cooperative-return-reports (Accessed: September 26, 2020).

Internal Revenue Service. 2024*e*. "SOI Tax Stats – Corporation Source Book (Publication 1053)." , https://www.irs.gov/statistics/soi-tax-stats-corporation-source-book-publication-1053 (Accessed: September 26, 2020).

Kwon, Spencer Yongwook, Yueran Ma, and Kaspar Zimmermann. 2024. "Data and Code for 100 Years of Rising Corporate Concentration." , https://doi.org/10.3886/E197985V1.

Lamoreaux, Naomi R. 2006. "Active Proprietorships, Partnerships, and Corporations – Entities, Receipts, and Profits." In *Historical Statistics of the United States, Earliest Times to the Present: Millennial Edition.* , ed. Susan B. Cater, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch and Gavin Wright. Cambridge University Press.

Means, Gardiner Coit, Donald Ewan Montgomery, John Maurice Clark, Alvin Harvey Hansen, and Mordecai Ezekiel. 1939. *The Structure of the American Economy.* Vol. 1, US Government Printing Office.

Mills, Lillian F., Kaye J. Newberry, and William B Trautman. 2002. "Trends in Book-Tax Income and Balance Sheet Differences." Working Paper.

National Archives and Records Administration. 2024. "Corporation Source Book Study File." , https://catalog.archives.gov/search-within/646618?levelOfDescription=fileUnit&limit=20&sort=naId%3Aasc (Accessed: May 3, 2024).

Navin, Thomas R. 1970. "The 500 Largest American Industrials in 1917." *Business History Review*, 44(3): 360–386.

Piketty, Thomas, and Emmanuel Saez. 2003. "Income Inequality in the United States, 1913–1998." *Quarterly Journal of Economics*, 118(1): 1–41.

Streuling, Guenther Fred. 1971. *A Comparison of Income Tax Regulations for Consolidated Returns and Generally Accepted Consolidation Practices for Financial Statements.*

U.S. Census Bureau. 1997. "1997 Economic Census: Manufacturing Reports." U.S. Department of Commerce Government Report. (Accessed: June 28, 2023).

U.S. Census Bureau. 2002. "2002 Economic Census: Manufacturing Reports." U.S. Department of Commerce Government Report. (Accessed: June 28, 2023).

U.S. Census Bureau. 2007. "2007 Economic Census Manufacturing Reports." , https://www.census.gov/data/tables/2007/econ/census/manufacturing-reports.html (Accessed: June 28, 2023).

U.S. Census Bureau. 2012. "2012 Economic Census Manufacturing Reports." , https://www.census.gov/data/tables/2012/econ/census/manufacturing-reports.html (Accessed: June 28, 2023).