NBER WORKING PAPER SERIES

SUPERSTAR FIRMS THROUGH THE GENERATIONS

Yueran Ma
Benjamin Pugsley
Haomin Qin
Kaspar Zimmermann

Working Paper 34194
http://www.nber.org/papers/w34194

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2025

We are grateful to Fatin Alia Md Ali, Chun Hei Hung, Nurul Amin Khan, Koichi Onogi, Sheena Flor Samonte, Ramraj Harikanth Sundararaj, Sue Cat Tuong, Zhiyuan Xu, Yingxuan Yang, Raymon Yue, and Zi Ye for excellent research assistance. We thank George Alessandria, Anmol Bhandari, Serguey Braguinsky, Eric Budish, Paco Buera, Cheng Chen, Gian Luca Clementi, Carter Davis, Lars Peter Hansen, Steve Kaplan, Stephen Redding, Esteban Rossi-Hansberg, Andrei Shleifer, Chad Syverson, Nico Trachter, Rob Vishny, Michael Weisbach as well as seminar and conference participants at Bank of Canada, BLS, Blue Collar Working Group, Boston College, BSE Summer Forum, U.S. Census Bureau, Chicago Booth, Chicago Fed, Clemson, Cleveland Fed, Colorado Boulder, Frankfurt Macro Seminar, HKUST, LAEF Labor Markets and Macroeconomic Outcomes Conference, MadMac Annual Conference, McGill, Minnesota Corporate Finance Conference, NBER Summer Institute, Notre Dame, Nova, Penn State, Philadelphia Fed, PUC, Purdue, Queensland Corporate Finance Conference, SAIF, SED, St. Louis Fed, Rice University, Rochester, UCSB Firm Dynamics and Growth Conference, and Western University, for helpful suggestions. We acknowledge financial support from the Macro Finance Research Program and the NSF CAREER Award. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1731 (CBDRB-FY25-P1731-R12291). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Superstar Firms through the Generations
Yueran Ma, Benjamin Pugsley, Haomin Qin, and Kaspar Zimmermann
NBER Working Paper No. 34194
August 2025
JEL No. D2, E2, L1, M1

ABSTRACT

We present new facts about the largest American companies over the past century. In manufacturing, top firms in the 1910s, 1950s, and 2010s predominantly date back to around 1900. Even as this special generation persists, turnover among top firms has been substantial. In contrast, in retail and wholesale, we do not observe a special generation among top firms. We show in a model of firm dynamics that a special generation can arise from an industrial revolution, through the adoption of a scalable technology and learning-by-doing. Top firm turnover is matched by standard idiosyncratic productivity shocks. Time-varying market size growth rates or entry costs are not sufficient to explain the facts. Among retailers and wholesalers, learning appears absent, so a special generation would be harder to sustain. Our results highlight the potential for lasting nonstationarity among the dynamics of top firms, which can result from the long echoes of technological change.

Yueran Ma
University of Chicago
Booth School of Business
and NBER
Yueran.Ma@chicagobooth.edu

Benjamin Pugsley
University of Notre Dame
Department of Economics
bpugsley@nd.edu

Haomin Qin
London Business School
Department of Finance
hqin@london.edu

Kaspar Zimmermann
Kiel Institute for the World Economy
and University of Hamburg
Kaspar.Zimmermann@ifw-kiel.de

1 Introduction

Production in the modern economy is highly skewed, with a small set of top companies—sometimes referred to as "superstar firms"—accounting for a considerable share of output (Berle and Means, 1932; Gabaix, 2011; Crouzet and Mehrotra, 2020; Autor et al., 2020; Braguinsky et al., 2024; Kwon, Ma, and Zimmermann, 2024).[1] Are new superstars born at the same pace over time? Is the dominance of existing superstars persistent? These questions attract much attention (The Economist, 2023; Draghi, 2024), but the answers are not yet fully understood.

In this paper, we collect new data and provide an extensive analysis of the largest American companies over the past century. In manufacturing, we find that the emergence of superstars is highly uneven over time—with a special generation from 1880 to 1920 remaining dominant among the largest manufacturers in the 1910s, 1950s, and 2010s—but top firms still experience substantial turnover. In contrast, in retail and wholesale, we observe that the emergence of superstars appears stable over time, without any special generation. We show in a model that a special generation among top firms can arise from the firm dynamics following an industrial revolution. We extend the Hopenhayn (1992) framework to allow for technology adoption and learning-by-doing, elements in the spirit of Chandler (1994), Klepper (1996), and Jovanovic and Rousseau (2005). In this case, a shock that reduces adoption costs of a more scalable technology, combined with modest first-mover advantages from learning, can match the special generation among top manufacturers; meanwhile, Hopenhayn (1992) idiosyncratic productivity fluctuations generate churning among top companies similar to what we observe in the data. Among retailers and wholesalers, we find that learning appears absent, which would make a special generation more difficult to sustain. Overall, we highlight the potential for lasting nonstationarity among the dynamics of top firms, resulting from the long echoes of technological change.

To trace out the identity of the largest American companies, we start with the *Fortune* list in 2018 (a recent example year before COVID), which covers the largest 1,000 companies by sales across all sectors. We also analyze the first *Fortune* list in 1955, which covers the largest 500 industrial (i.e., manufacturing and mining) companies by sales. Before 1995, *Fortune* published additional lists of top companies in other sectors on a more ad hoc basis, and top

[1] For example, total sales of the top 1,000 firms in the 2018 *Fortune* list we study exceed 40% of U.S. private sector gross output. We use the term "superstar firms" as a simple reference for these large companies that have achieved an extraordinary size relative to other businesses. We do not imply any impact on welfare.

retailers/wholesalers have received relatively consistent coverage. In addition, we obtain a list of the largest 500 industrials by assets in 1917 compiled by Navin (1970). For each company, we research its origin using information from company sources (e.g., websites, filings, and anniversary documents) whenever possible, and business history websites otherwise.

Among top industrial firms in 2018, 1955, and 1917, we find that birth years always cluster around 1880 to 1920. We refer to this cohort as a "special generation." Because of its lasting influence, the largest industrials have become much older: for example, their median age was around 30 in 1917, 60 in 1955, and 100 by 2018. The patterns are similar if we drop companies with complex histories (e.g., major mergers or spinoffs). Despite the persistence of the special generation, the persistence of individual companies is not high. Through a detailed comparison, we find that 21% of the top industrials on the 1955 *Fortune* list remain on the 2018 *Fortune* list.[2] Correspondingly, the 2018 *Fortune* list has many "late bloomers": among the 388 industrials in this list, 137 were born between 1880 and 1920, but only 50 were among the top 388 industrials in 1955; the rest (and the majority) is represented by "late bloomers."

Is the special generation of top industrials unique to the U.S.? From Chandler (1994), we obtain the largest 200 German industrials by assets in 1913 and 1953, along with the largest 200 British industrials by market capitalization in 1919 and 1948. We supplement these lists with data on the largest German and British industrials in 2018. The patterns in Germany turn out to be similar: top industrials were young in the early 20th century, and fairly old by 2018; the cohort from around 1900 continues to dominate. In contrast, top British industrials are relatively young today, without a special generation. Curiously, these findings align with Chandler's remark that the German experience in the Second Industrial Revolution "is closer to that of the United States than to that of Britain." In his view, "in Germany as in the United States, but much more than in Britain, entrepreneurs did make the investment in production facilities and personnel large enough to exploit the economies of scale and scope" (Chandler, 1994). The international evidence also suggests that the special generation among top industrials is not just a result of country-specific regulations (given the similarities in the U.S. and Germany), or military buildup in the world wars (which would be relevant in the U.K. too).

We then turn to retail and wholesale. Today, the birth years of top retailers and wholesalers cluster around 1960 to 1980. In the 1950s, however, top companies primarily date back to

[2]Another 21% "survived" on the 2018 *Fortune* list through acquirers, although this may overstate persistence as some firms were acquired following operating weakness (e.g., Union Carbide & Carbon was acquired by Dow Chemical after accidents and poor performance, National Steel Corporation was sold in bankruptcy to U.S. Steel).

around 1900. After collecting additional data from 1970 and 1995, we confirm that the largest retailers and wholesalers have always stayed 60 to 70 years old on average, without a special generation. The stability of the age distribution among large retailers and wholesalers presents an interesting contrast with the results among large industrials.

For top services companies, today their birth years cluster around 1960 to 1980, and few large ones existed before then. The business size distribution data in Kwon, Ma, and Zimmermann (2024) also indicate that few services companies would qualify for the largest businesses in the economy until recently. In 2018, the large services in *Fortune* 1,000 are fairly young, with a median age of 43, which resembles the youth of top industrials in 1917. It is possible that the cohort of services firms born around the Third Industrial Revolution can form a special generation, like the cohort of industrial firms born around the Second Industrial Revolution, but we need several more decades to verify whether this is the case (e.g., the median age of the largest services firms in 1995 *Fortune* 1,000 list was 31, and it is difficult to ascertain whether median age changing by 12 years over 23 years represents a deviation from stationarity).[3]

We investigate firm dynamics models to understand the economic forces that can shape the empirical evidence we observe. We start with the canonical Hopenhayn (1992) model. In this benchmark model, top firms emerge from the combination of large permanent advantages in productivity drawn at entry *and* persistent positive productivity shocks accumulated over their lifetime. The permanent and persistent components of productivity generate a stationary hump-shaped age distribution among top firms. Very young firms are less represented as they have not had enough time to grow large, and very old firms are less represented as they experience more attrition with age. Accordingly, we cannot detect a special generation from a single cross section of top firms. We need multiple cross sections to see that birth years cluster around the same time period.

The benchmark model is stationary by design, so it does not directly match the evidence among top industrials, where the cohort from around 1900 enjoys lasting dominance for many decades. We then extend it to examine three hypotheses about the emergence of the special generation among top industrials. First, we follow Chandler (1994) and consider the adoption of a capital intensive modern technology that increases the returns to scale. The adoption cost of the modern technology falls during the Second Industrial Revolution between 1860 and 1900

[3]For other sectors, the large companies in finance as well as transportation, communications, and utilities are old on average, but their birth years do not display much clustering. We do not focus on these sectors because they can be heavily influenced by regulation, and historical data on large firms are less consistently available.

(Gordon, 2016), and entrants can choose the technology they use (whereas incumbents cannot due to organizational rigidity). We also allow firms to become more efficient over time through learning-by-doing (e.g., accumulation of organizational capital, know-how, or customer base), which is an important mechanism for first-mover advantages in the view of Chandler (1994). We calibrate the model by matching a set of moments along the transition path resulting from a perfect foresight shock that reduces the modern technology adoption cost, starting from an initial stationary equilibrium with high adoption costs and no modern technology use. We use data moments that capture overall firm dynamics around 2018 (e.g., growth rates and exit rates by age and entrant size distribution), together with moments about top firms in 2018 and 1955 along the transition path (top firms' median age, revenue share, and fraction born between 1880 and 1920). The calibration obtains results that align with the empirical evidence among the largest American industrials. Rapid declines of the adoption cost of the modern technology around the Second Industrial Revolution, combined with a modest amount of learning, can give rise to the special generation among top industrials. With common levels of volatility for firm-level idiosyncratic shocks, individual companies at the top still change substantially over time even as the special generation persists. Finally, if the adoption of the modern technology happens slowly, as Chandler (1994) hypothesized about the U.K., then a special generation may not take hold.

Second, we consider time-varying entry dynamics due to high market size growth for a period of time, which may bring in a larger number of entrants. In the data, U.S. manufacturing experienced high growth until 1970 and low growth afterwards. Accordingly, we build on the baseline model to examine the nonstationary equilibrium resulting from a perfect foresight shock to market size growth between 1860 and 1970. We maintain homogeneous technology as in the baseline model. We continue to allow for learning, which may strengthen the staying power of important cohorts; doing so also enhances comparability with the previous calibration. We calibrate this model to the same set of data moments, replacing model parameters that govern heterogeneous technology with those that govern time-varying market size growth rates. In this case, we find that the most prevalent cohort among top firms in 2018 comes from the 1960s, shortly before the high growth period ends: this cohort enjoys the large size and has had less subsequent displacement compared to their predecessors. Overall, the timing of the special generation (1880 to 1920) and the timing of high manufacturing growth (lasting until 1960s) do not align.

Third, we consider time-varying entry costs, which are commonly used to capture the effects of regulatory changes. In particular, rising entry costs may diminish the size of later entering cohorts. We build on the baseline model as before, and examine the nonstationary equilibrium resulting from a perfect foresight shock that raises entry costs after 1860. Again, we maintain homogeneous technology as in the baseline model, and continue to allow for learning to enhance comparability with previous calibrations. We calibrate this model to the same set of data moments, now with model parameters that govern time-varying entry costs (instead of time-varying technology adoption costs or market size growth rates). To generate a special generation of top firms born around 1900, the rise of entry costs needs to accelerate starting in the early 1900s. However, a key problem is such increases in entry costs per se do not make firms born around 1900 stand out against those born earlier in the 1800s. Producing a special generation concentrated around 1900 then requires fast attrition of older firms, for example through high volatility of productivity and high fixed costs. Doing so results in too much churn among top firms and no overlap between top firms in 1955 and 2018.

Finally, we use the model to investigate why top retailers and wholesalers do not feature a special generation. By calibrating the baseline model to the data on retail and wholesale trade, we find that learning appears absent in this sector. In particular, entrant size already displays a large degree of heterogeneity, so the presence of learning would make high productivity firms last too long and top firms would be too old relative to the data. Without learning, even if firms in a given cohort obtain technological or other advantages, subsequent entrants can displace them more easily. Since retail operations are spread out among many establishments (Oberfield et al., 2024), it is possible that achieving efficiency gains throughout the organization is challenging. It is also possible that imitation is easier in retail, making it more difficult to accumulate efficiency gains within the firm.

Taken together, the emergence of superstar firms may cluster in time if a cohort has a particularly strong edge relative to both firms that came before and potential entrants afterwards. In our model of technological heterogeneity, declining adoption costs of the scalable technology generate the advantage of the special cohort relative to firms that came before; the accumulation of productivity over time through learning gives them first-mover advantages relative to potential entrants afterwards. In this case, prospective superstar firms are not born at a constant pace, and the persistence of the special generation does not necessarily imply staleness or lack of dynamism among top firms.

Literature Review Our work relates to three sets of literature. First, we add to the firm dynamics literature. Many papers provide insights about the population of businesses, whereas few study the dynamics of top firms. Influential work by Luttmer (2011) highlights the importance of growth rate heterogeneity in explaining the relatively young age of the largest American companies. Guntin and Kochen (2025) analyze how top firms grow in 30 years of Spanish micro data, and emphasize technological heterogeneity as a key element. Both papers focus on the stationary equilibrium. We examine the evolution of top firms over a long period of time: we uncover the lasting influence of special cohorts—and with it the aging of large industrials over time—combined with the fluidity of individual companies at the top. We investigate how firm heterogeneity matters not just in the stationary equilibrium but also in response to shocks over time, leading to the century-long dominance of the special generation among top firms.[4] Our findings also complement work on slow transitional dynamics (Atkeson and Kehoe, 2007; Buera and Shin, 2013).

Among firm dynamics studies, two strands of work relate to cohort effects. One strand is models of shakeout (Jovanovic and MacDonald, 1994; Klepper, 1996; Beraja and Buera, 2024), which can lead to persistent dominance of firms entering around the time of key innovations. Classic shakeout studies scrutinize specific markets. Our evidence suggests that the manufacturing industry as a whole has seen a special generation among its top firms for over a century, which may be viewed as a shakeout wave at a large scale. Another strand is analyses of cohort effects at the business cycle frequency (Moreira, 2017; Sedláček and Sterk, 2017), which find that companies born in recessions remain smaller than those born in booms. These studies and others (Clementi and Palazzo, 2016; Lee and Mukoyama, 2018) show the importance of entry and exit in the propagation of aggregate shocks. We focus on slow-moving, century-long cohort effects resulting from the Second Industrial Revolution.

Second, researchers and the general public have always been interested in the large companies in the economy (Marshall, 1920; Berle and Means, 1932; Collins and Preston, 1961; Hannah, 1976; Prais and Reid, 1976; Autor et al., 2020), and our analyses add new findings to this inquiry. To our knowledge, two previous articles tabulated the overall age of *Fortune* 500 companies for 1995 (Harris, 1996) and 2009 (Stangler, 2009). Luttmer (2011) examined the average age of companies with more than 10,000 employees in 2008. Several articles have

[4]Our findings have implications for dynamism of top firms (of all ages) over the long run. Decker et al. (2016) uncover declines in dynamism among *young* firms since the 2000s, which could be attributed to increasing factor adjustment costs (Decker et al., 2020).

noted constant churn among *Fortune* or similar lists (Hannah, 1998; Louçã and Mendonça, 2002; Stangler and Arbesman, 2012). We show that the persistence of the special generation among top firms and the displacement of individual companies coexist. We also document that strong cohort effects occur in some industries but not in others, as well as the similarity between American and German industrials, which suggest that these patterns do not just result from economy-wide regulatory policies in the U.S.

Finally, the literature on innovation has also been interested in entrants versus incumbents (Schumpeter, 1942; Aghion and Howitt, 1992; Aghion and Tirole, 1994; Klette and Kortum, 2004; Acemoglu et al., 2018). Garcia-Macia, Hsieh, and Klenow (2019) suggest that most growth comes from innovation and improvement by incumbent firms, using U.S. Census data from 1983 to 2013. To the extent that some generations of new firms are more powerful than others, the early 20th century could provide different results for such an analysis.

In the rest of this paper, Section 2 describes the data collection; Section 3 presents the empirical facts; Sections 4, 5, and 6 examine firm dynamics models; and Section 7 concludes.

2 Data

We begin with the list of the largest American companies by sales revenue published by *Fortune* since 1955. After 1995, *Fortune* has been publishing an annual list of the largest 1,000 companies by sales combining all industries. Before 1995, *Fortune* published an annual list of the largest industrial (i.e., manufacturing and mining) companies by sales, together with additional lists such as the 50 largest companies in other industries. Tabulations of the largest firms are most consistently available for industrials and retail/wholesale, so our analyses focus on these sectors.[5] In addition to the *Fortune* lists, we study the 500 largest American industrials in 1917 compiled by Navin (1970), which is based on firm size by assets due to data availability.[6] We use these lists because it is otherwise difficult to obtain comprehensive information about the identities of the largest companies, especially for the earlier decades.

We then collect information about the origin of these top companies, focusing on 2018

[5] Historical *Fortune* lists available online (e.g., via CNN Money) do not necessarily display the actual name of the company at the time of the list, but use the name of their successors (e.g., Esmark for Swift in the 1955 list). We recover the actual name of the company at the time of each *Fortune* list from the original print publications.

[6] "I, too, would have used sales had I been able, but very few companies reported their sales in 1917 whereas many companies reported their assets, and most reported their capitalization from which their assets could be estimated." (Navin, 1970)

(as an example year before COVID), 1955 (the first year of *Fortune* lists), and 1917 (the earliest comprehensive list we can find). Whenever possible, we use information from company sources including websites, SEC filings (e.g., annual reports), and anniversary documents. Otherwise, we use business history websites and Wikipedia.[7] If firm A acquired firm B, we use the origin of firm A. If two or more firms joined as equals, we use the earliest origin among them. In the rare cases where a firm was formed as a subsidiary and later spun off, we prioritize using the subsidiary's formation year. Overall, we aim to use the beginning of business operations (i.e., entry of the enterprise into production activities), and mimic the notion of entry in the firm dynamics literature. Our results are similar if we drop cases where the origin year is hard to determine (e.g., due to spinoffs, major mergers, or consolidating a large number of small predecessors). For the *Fortune* list in 1995, Harris (1996) tabulates the founding years of the top 500 companies, which are exactly the same as founding years from our research for 67% of the companies and within ten years for 82% of the companies. We also compare with firm age distribution among the largest employers in the census Business Dynamics Statistics (BDS) dataset in Section 3.1.1, but firm age in BDS data is left censored for those formed before 1976.

We do not use off-the-shelf data on incorporation year (e.g., from Dun & Bradstreet), which may not accurately represent the economic history of a company. First, firms sometimes reincorporate for legal reasons (e.g., when they move states or merge), so age based on the incorporation year may not reflect the actual age. Second, the incorporation year omits the firm's early origins before incorporation. As a result, off-the-shelf data on incorporation year tend to contain many large firms that appear excessively young. We do not focus on size by stock market capitalization given the time-varying selection into publicly listed firms. For instance, the number of public firms has decreased by one half between the 1990s and the 2010s, and public firms have become less representative of the economy (Doidge, Karolyi, and Stulz, 2017; Schlingemann and Stulz, 2022). Large firms staying private were also common in the early 20th century (e.g., Ford was private until the 1950s). In addition, in the early decades, compiling comprehensive data on firms' market capitalization is challenging given the multitude of trading venues (Kuvshinov and Zimmermann, 2022); for example, CRSP covers only NYSE firms until the 1960s. Furthermore, market capitalization can be affected by risk premia or biased beliefs (Fama and French, 1992; Bordalo et al., 2024).

[7] For the 2018 *Fortune* 1,000 list, we can use data from company sources for 830 firms. For the 1955 *Fortune* 500 list, we can use data from company sources for 216 firms (e.g., many companies did not build a website before they ceased to exist).

3 Empirical Facts

This section presents empirical facts about the largest American companies at different points in time. We focus on birth years and the age distribution of top companies, as well as the degree of persistence among them, since these analyses are informative for firm dynamics research and can be performed reasonably reliably. Tracing out the growth trajectories of individual companies is difficult to implement, as limited financial information exists for the early years of most companies (including the many large companies that are fairly old).

3.1 Industrials

We start with the industrial sector (manufacturing and mining), where we have the most extensive data about top firms over time.

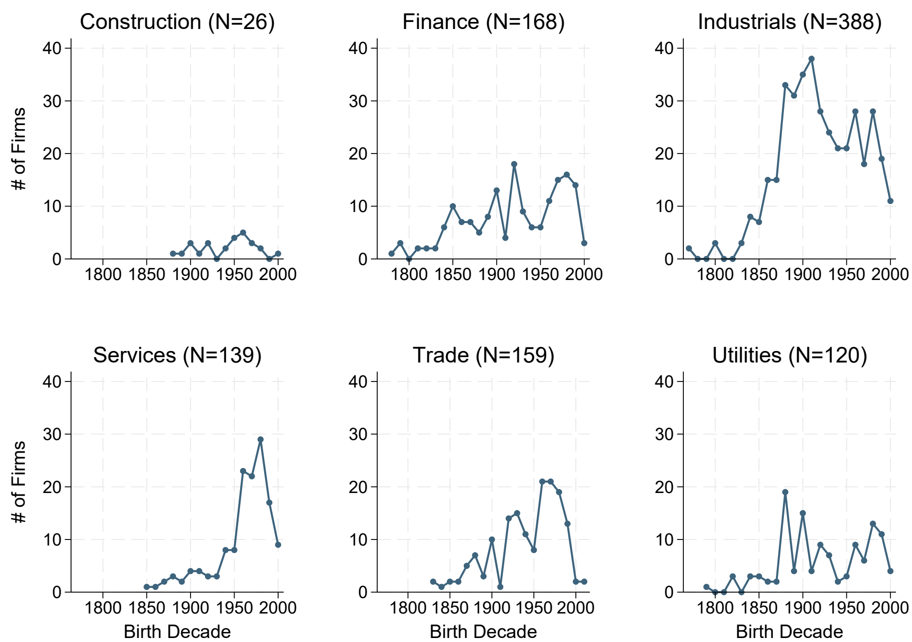
3.1.1 Birth Years of Top Firms

Top Firms in 2018 In Figure 1, we plot the number of 2018 *Fortune* 1,000 companies that originated in each decade, based on the firm's industry in 2018. For industrials, we observe substantial clustering of birth years around the turn of the 20th century: the cohort from 1880 to 1920 is the most represented among today's largest industrials. We verify below that this cohort is a "special generation," using data on top industrials in 1955 and 1917. We discuss other sectors in Section 3.2. Figures IA1 and IA2 show that the results are robust to alternative ways to weight the birth decade distribution (e.g., by sales or employment), so we will focus on distributions by number for parsimony. Figure IA3 shows that the results are also similar if we include more private firms using the *Forbes: America's Largest Private Companies* list for 2018, since the *Fortune* list excludes private companies that do not file financial statements with a government agency. In Figure IA3, we combine the *Fortune* and *Forbes* lists and take the largest 1,000 by sales, which contain 171 companies from the *Forbes* list not in the *Fortune* list.

In Figure IA4, we check that the results are similar if we restrict to companies with a clear main lineage: we drop firms with complex histories, including those that involve mergers of equals (which complicate the main lineage), and those that are spin-offs where we do not know the origin year of the spun-off entity. The birth year distribution in this case (solid line with circles) is almost identical to that in our baseline (solid line with triangles). Finally, although Census data do not have the exact age of firms formed before 1976, in Figure IA5 we draw a comparison based on broad age groups with firms having more than 10,000 employees in the

Figure 1: Birth Years of 2018 *Fortune* 1,000 Companies

This figure shows the number of birth years per decade for the 2018 *Fortune* 1,000 companies. Companies are assigned to main sectors based on their industries in 2018. The main sectors correspond to SIC codes 15-17 (construction), 10-14 and 20-39 (industrials), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services); N denotes the number of firms in each main sector.



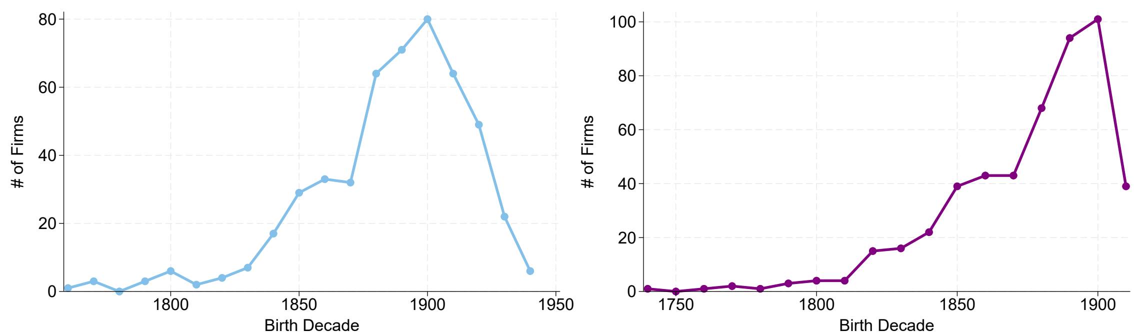
Business Dynamics Statistics (BDS) dataset, which contain 1,196 firms in 2018. In this case, 67% were born before 1976 (gray bars), compared to 75% in *Fortune* 1,000 firms (blue bars). The distributions based on these coarse age groups align well between our data and BDS data.[8]

Top Firms in 1955 In 1955, *Fortune* published the first list of the largest 500 industrials by sales. Panel A of Figure 2 plots the birth years of the largest 500 industrials in the 1955 *Fortune* 500 list. Interestingly, in this case we still observe clustering around the turn of the 20th century. Accordingly, the age distribution of the largest industrials in 1955 is very different from that in 2018: the median age of the largest industrials in 1955 and 2018 is 57 and 98, respectively. Figure IA6 provides further visualization. The solid line with triangles shows the birth year distribution of the largest 388 industrials in 2018. The dashed line with circles shows what it

[8]Another possible question is whether firms at birth belonged to the same sectors as when we observe them in the top firm lists. We collect information on firms' businesses at the time of founding. Among the *Fortune* 2018 companies, 87% belong to the same main sector. Results are similar for the largest companies in other years. Therefore, the patterns above hold if we instead use firms' industry group at birth. The stability of the business over time is consistent with findings in (Kaplan, Sensoy, and Strömberg, 2009).

Figure 2: Birth Years of the Largest 500 Industrials in 1955 and 1917

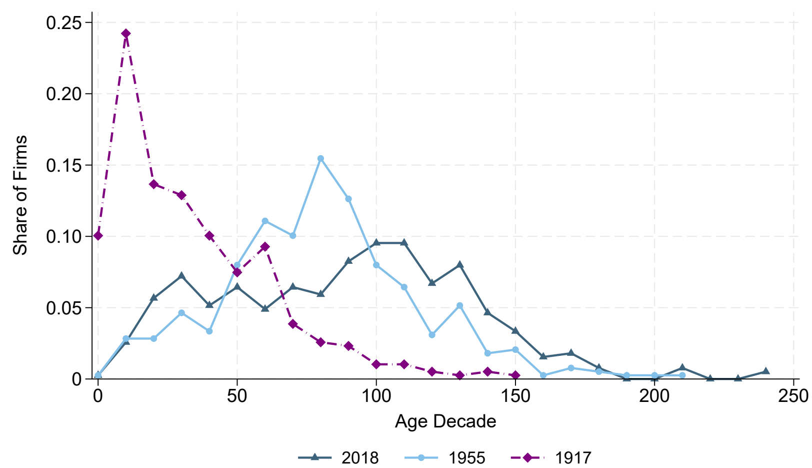
Panel A shows the number of birth years per decade for the 1955 *Fortune* 500 industrials (by sales). Panel B shows the number of birth years per decade for the largest 500 industrials (by assets) in 1917 from Navin (1970).



Panel A. Birth Years in 1955 *Fortune* List   Panel B. Birth Years in 1917 Navin (1970) List

Figure 3: Age Distribution of the Largest Industrials

The solid line with triangles shows the age decade distribution of the largest 388 industrials (by sales) in the 2018 *Fortune* list (which contains 388 industrial firms). The solid line with circles shows the age decade distribution of the largest 388 industrials (by sales) in the 1955 *Fortune* list. The dashed line with diamonds shows the age decade distribution of the largest 388 industrials (by assets) in 1917 from Navin (1970).



would look like if the age distribution of top firms were stationary, by shifting the birth year distribution in 1955 (the solid line with circles) by six decades. The actual 2018 distribution is very different from this benchmark, and instead has more overlap with the 1955 distribution.

Top Firms in 1917 Before the *Fortune* lists started in 1955, we have one additional tabulation of the largest 500 industrial companies by assets in 1917 from Navin (1970). Panel B of Figure 2 plots the birth years of these largest industrials in 1917. Remarkably, we still see the clustering of birth years around 1900 for the largest industrials in 1917. At this time, the median large industrial is only 30 years old, compared to 57 years old in the 1955 *Fortune* list and 98 years old in the 2018 *Fortune* list. Figure 3 shows the age distribution of top industrials across these three points in time, which is evidently nonstationary.

3.1.2 Persistence of Top Firms

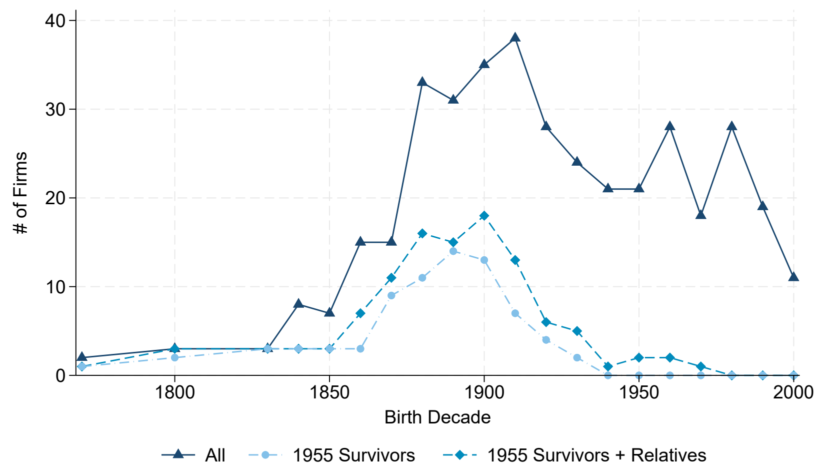
Given the cohort from around 1900 has been dominant for a very long time, a natural question is whether the largest firms over time are the same companies. Indeed, the prevalence of old firms among the recent *Fortune* lists has led some observers to postulate that corporate giants are hard to topple (The Economist, 2023). We analyze the persistence of top firms in the following, and perform a detailed comparison of the largest industrials in the 1955 *Fortune* list and the 2018 *Fortune* list. Interestingly, even though the cohort born around 1900 has remained prominent among the largest industrials for a century, we find that the exact companies have changed substantially.

We compare the 1955 *Fortune* list and the 2018 *Fortune* list in several ways. Since there are 388 industrial firms in the 2018 *Fortune* list, we focus on the top 388 industrial firms in the 1955 *Fortune* list. First, we look at the fraction of the companies on the 1955 *Fortune* list that remain on the 2018 *Fortune* list. Overall, 21% of the companies on the 1955 *Fortune* list remain on the 2018 list. Table IA1 presents more detailed information for firms born in each decade. Columns (1) to (4) show the number of firms born in each decade for the top 388 firms in the 1955 list and the 2018 list, respectively. Columns (5) and (6) show the number of top firms in the 1955 list that survived on the 2018 list, for firms born in each decade. For example, 37 of the top firms in the 1955 list were born in the 1910s, only nine of which survived to the 2018 list, yet a total of 38 top firms in the 2018 list were born in that decade.

Second, we recognize that some of the companies on the 1955 *Fortune* list were subsequently acquired (e.g., Quaker Oats was acquired by PepsiCo). In this case, another 21% of the companies on the 1955 *Fortune* list "survived" on the 2018 *Fortune* list through their acquirers, so 42% in total either "survived" on the 2018 *Fortune* list as either the main entity or through their acquirers. Columns (7) and (8) of Table IA1 shows these direct plus indirect survivors by birth decade. Finally, another 22% of the top firms in the 1955 list were acquired by foreign firms (which do not qualify for the 2018 *Fortune* list). "Survival" through acquirers may overstate the persistence since some firms were acquired following operating weakness (e.g., Union Carbide & Carbon was acquired by Dow Chemical following accidents and poor performance, National Steel

Figure 4: Overlap between the Largest Industrials in 2018 and 1955

This figure shows the overlap between the largest 388 industrials in 2018 and 1955. The solid line with triangles plots the number of companies founded each decade among the largest 388 industrials in 2018. The dashed line with circles represents the number of birth years per decade for firms on the 2018 list that were also among the largest 388 industrials in 1955. The dashed line with diamonds represents the number of birth years per decade for firms on the 2018 list that were either among the largest 388 industrials in 1955 or have relatives among the largest 388 industrials in 1955 (e.g., through a spinoff or acquisition).



Corporation was sold in bankruptcy to U.S. Steel).

Figure 4 plots the birth year distribution of the largest industrials in 2018 that were also among the largest 388 in 1955, either directly or through relatives (i.e., acquired a firm among the largest 388 industrials in 1955 or is a spinoff of a 1955 top firm). The birth year distribution for all top industrials in the 2018 is also shown. Among the largest industrials today which were born around 1900, many were not at the top in 1955. Some examples include Clorox, Coty, Cummins, Harley-Davidson, Harris, Lear, Nucor, Parker-Hannifin, VF, and Xerox. For more examples, Table IA2 shows the top ten firms in food and metal manufacturing in 1955 and 2018: the birth years are remarkably similar, but the exact companies are rather different.

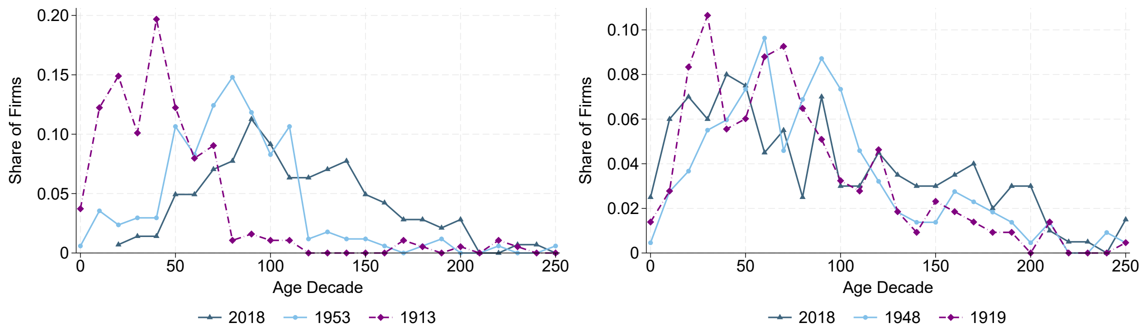
3.1.3 Germany and United Kingdom

A common question is whether similar cohort effects exist among top firms in other countries. From Chandler (1994), we are able to obtain the list of the largest 200 German industrial firms by assets in 1913 and 1953, and the largest 200 British industrial firms by market value of shares in 1919 and 1948.[9] We supplement these historical lists with information about the

[9]The historical British lists rely on size by market capitalization of shares (and therefore restrict to public firms) due to sparse data on sales or assets (Prais, 1976; Meeks and Whittington, 2023).

Figure 5: Age Distribution of the Largest German and British Industrials

Panel A shows the age decade distribution of the largest 200 German industrials by sales in 2018 (solid line with triangles), and by assets in 1953 (solid line with circles) and 1913 (dashed line with diamonds). Panel B shows the age decade distribution of the largest 200 British industrials by sales in 2018 (solid line with triangles), and by market capitalization in 1948 (solid line with circles) and 1919 (dashed line with diamonds).



Panel A. German Industrials

Panel B. British Industrials

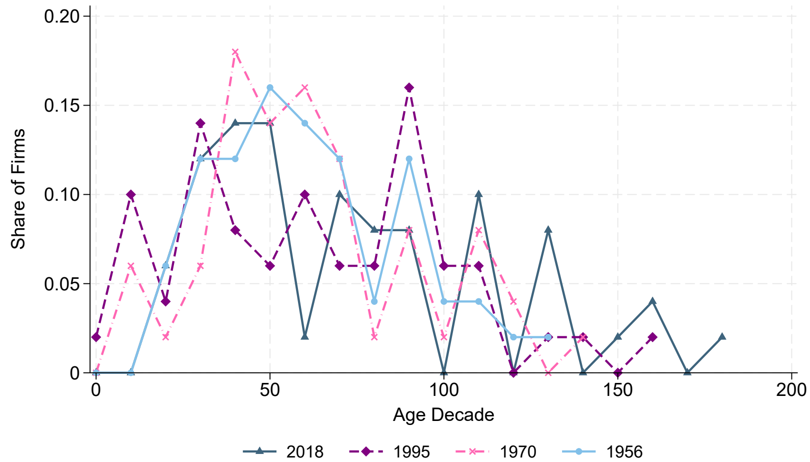
largest firms in 2018. For Germany, the newspaper *Die Welt* published a list of the largest 500 German firms by sales in 2018. The list is based on annual reports or information provided by the companies at the request of the editors. We select the 225 industrials on that list and drop those that are subsidiaries of other firms (e.g., Audi). We do not use ORBIS data for Germany because of poor coverage (Bajgar et al., 2020), and many large firms are missing in ORBIS. For the U.K., we get the largest 200 industrials by sales in 2018 using ORBIS data (ORBIS coverage is better for the U.K., and we are not aware of good alternatives). We then research the history of these companies, following the same approach for U.S. firms.

Figure 5 shows the age distributions for the largest German (Panel A) and British (Panel B) industrials in the 1910s, mid 1900s, and 2018. The patterns among the largest German industrials are similar to those among the largest American industrials (e.g., Figure 3). The modal cohort comes from around 1900, and top firms were very young in the early 1900s but fairly old by 2018. Meanwhile, the largest British industrials behave quite differently. The age distribution looks rather stationary, and no particular cohort has lasting dominance.

In his study of the largest American, German, and British industrial companies, Chandler (1994) did not examine their birth years, but he remarked that the German experience through the Second Industrial Revolution "is closer to that of the United States than to that of Britain." In particular, "in Germany as in the United States, but much more than in Britain, entrepreneurs did make the investment in production facilities and personnel large enough to exploit the

Figure 6: Age Distribution of the Largest Retailers and Wholesalers

The solid line with triangles shows the age decade distribution of the largest 50 retailers and wholesalers in the 2018 *Fortune* list. The dashed line with diamonds shows the age decade distribution of the largest 50 in the 1995 *Fortune* list. The dashed line with crosses shows the age decade distribution of the largest 50 in 1970. The solid line with circles shows the age decade distribution of the largest 50 in 1956.



economies of scale and scope [...]," and Britain provides "a counterpart" where "fewer such firms appeared, and they grew in a slower and more evolutionary manner" (Chandler, 1994). Correspondingly, the cohort from the Second Industrial Revolution is more likely to have lasting staying power in the U.S. and Germany. Our data indeed reveal such patterns. The international evidence also suggests that the special generation among top industrials is not just a result of country-specific regulations (given the similarities in the U.S. and Germany), or military buildup due to the world wars (which would be relevant in the U.K. too).

3.2 Retail/Wholesale Trade and Other Sectors

We now turn to other sectors and draw comparisons with the findings among industrials.

Retail/Wholesale Trade Figure 1 above shows that the modal cohorts among top retailers and wholesalers in 2018 come from the 1960s to the 1980s. Before 1995, *Fortune* separately tabulated the largest 50 merchandising firms by sales, which first appeared in 1956. Figure IA7 shows that the largest retailers and wholesalers in the 1950s primarily date back to the late 1800s and early 1900s (e.g., grocery chains and department stores), which is rather different from the birth year distribution in the 2018 *Fortune* list. Figure 6 shows that the age distribution of the largest 50 retailers and wholesalers is similar in 1956 and 2018: their median age in 1956 and 2018 is 62 and 75, respectively. To further verify the stability of the age distribution of the largest retailers and wholesalers, Figure 6 also shows the age distribution of the largest 50 retailers and wholesalers in 1970 and 1995.[10] Correspondingly, the persistence of individual top companies is limited. Only nine companies among the top 50 retailers and wholesalers in 1955 remain among the 50 largest in 2018; another eight companies were acquired by Macy's, three by Albertsons, and three more by other top 50 retailers in 2018. Overall, the patterns among top retailers and wholesalers are quite different from those among top industrials, and later generations of large retailers and wholesalers have continuously displaced previous ones.

Services and other Sectors For services, in Figure 1 above, we observe clustering of birth years around the 1960s to the 1980s. Few large services existed in earlier decades. For example, in 1970, *Fortune* provided a one-off analysis of companies that had larger sales than the 500th industrial firm. They were able to find two firms in business services (Sperry & Hutchinson and Dun & Bradstreet), two hotels (Holiday Inns and Hilton Hotels), and five in entertainment (Kinney National Service, Music Corporation of America, United Artists, Columbia Pictures, and MGM).[11] We also validate the lack of large services earlier using tabulations of corporations by size of sales and industry in the Statistics of Income (SOI) published by the Internal Revenue Service (IRS) starting in 1959 (Kwon, Ma, and Zimmermann, 2024). We estimate the sales cutoff for the top 1,000 (and top 500) companies in the economy, and how many firms in services (as well as other industries) are above this threshold, shown in Figure IA8. Using 1959 as an example, the SOI data suggest that 19 firms (five firms) among the top 1,000 (top 500) by sales are in the services industry. For finance, some large companies date back to the 19th century (e.g., banks) whereas others originated in the second half of the 20th century (e.g., asset management firms and nonbanks). For utilities, mining, and construction, there is no obvious pattern but these industries also have fewer very large firms.

4 Baseline Model of Firm Dynamics

We examine firm dynamics models to investigate forces that can shape the landscape of top firms. In this section, we lay out the canonical Hopenhayn (1992) model of industry dynamics, which has a well-known stationary equilibrium and serves as a useful benchmark on which we will build. In Section 5, we investigate extensions of the model that can generate a special

[10] We use 1970 because after 1966 *Fortune* tabulated the largest 50 retailers and dropped the wholesalers, except for 1970 where the largest wholesalers are separately tabulated, and we use 1995 which is the first year *Fortune* tabulated the largest 1,000 companies combining all sectors.

[11] Given the scarcity of large services companies, *Fortune* did not provide much information about them until 1995 when the top 1,000 companies combining industrials and non-industrials were tabulated in one list.

generation among top firms similar to what we observe in the industrial sector. In Section 6, we analyze the stationary age distribution among top retailers and wholesalers. Additional computational details are provided in Internet Appendix IA2.

4.1 Setup

We model the industrial sector in partial equilibrium. There is a single good produced by many firms. Sector demand is given by a CES function $D(p_t) = L_t p_t^{-\epsilon}$, where ϵ is the demand elasticity and L_t is the market size with an exogenous growth rate η. Each firm i operates a decreasing returns to scale technology $y_{iat} = z_{iat} n_{iat}^{\theta}$ in labor, n_{iat}. We use subscript a to denote firm age and t for calendar time. Although firm decisions do not yet depend explicitly on age, we include firm age as a state variable to track its distribution. Firm i's productivity $z_{iat} = e_i s_{iat}$ has two components. As is common, the stochastic component, s_{iat}, evolves as an AR(1) process in logs over the firm's lifetime, i.e., $\log s_{iat} = \rho \log s_{ia-1t-1} + \sigma_\varepsilon \varepsilon_{it}$, with initial condition s_{i0} drawn at birth from distribution F. Here, ρ is the persistence of the stochastic component, σ_ε is its volatility, and ε_{it} is i.i.d. standard normal. We add a permanent component, e_i, which is drawn at birth from distribution G.

This setup models firms' life cycles as driven by a mix of ex ante and ex post heterogeneity (Sterk, Sedláček, and Pugsley, 2021). Log productivity for firm i of age a in period t can be written as the moving average:

$$\log z_{iat} = \log e_i + \rho^a \log s_{i0} + \sum_{k=0}^{a-1} \rho^k \sigma_\varepsilon \varepsilon_{it-k}. \tag{1}$$

The permanent type and initial condition terms in Equation (1) form an ex ante predictable life cycle. When the distribution of initial conditions, F, differs from the long-run distribution for $\log s_{it}$, as we will assume, the ex ante terms implicitly introduce age dependence to the stochastic component. As a firm ages, log productivity gravitates towards its long-run mean, $\log e_i$. Along the way, the realized life cycle may deviate from this ex ante path due to the accumulation of persistent ex post shocks captured by the last term.

Each period, operating firms pay a fixed overhead cost of c_f units of labor, and hire variable labor n_{iat} in a competitive labor market. There are no adjustment costs. They sell output in a competitive market at price p_t. Labor is the numeraire, so the wage $w_t = 1$. Firms discount the future using a constant discount rate r in *goods*, and face an exogenous probability of exit δ. They may also choose to exit endogenously and avoid paying any future fixed costs. Firms differ only in age and productivity. We next characterize recursively, for a given firm, its optimal labor choice and exit decision.

Recursive formulation Let $V_t\left(a, e_i, s_{iat}\right)$ be the value in period t of firm i with age a, permanent productivity component e_i, and stochastic component s_{iat}. Then V_t satisfies the following Bellman equation:

$$V_t = \max_{x\in\{0,1\}} \left\{ \max_n \left\{ p_t z_{iat} n^\theta - n \right\} - c_f + \frac{1-\delta}{1+r} \frac{p_t}{p_{t+1}} E_t \left[V_{t+1}\left(a+1, e_i, s_{i,a+1,t+1}\right)\right], 0 \right\}, \tag{2}$$

with productivity $z_{iat} = e_i s_{iat}$. Knowing current productivity, firms choose to produce and continue, or choose to exit. The exit decision is forward looking, and depends on the components of productivity. We let $x_t(a, e, s)$ be the (indicator) exit policy function that solves Equation (2). The continuation value also incorporates the probability of exogenous exit δ, in addition to discounting at the real rate r.[12] For any period t, we let $n_t\left(z\right)$ be the labor demand function and $y_t(z)$ the implied output. As static decisions, these depend only on the level of productivity z_{iat} and price p_t. The policy functions depend on calendar time t only through the path of p_t, and we maintain the time subscripts in anticipation of characterizing transition paths in Section 5.

Potential entrants and free entry Potential entrants evaluate the expected value of entering as an age zero firm against entry costs. We assume there is a perfectly elastic supply of entrants when this expected value equals a one-time sunk entry cost of c_e units of labor, i.e., the free entry condition:[13]

$$E\left[V_t(0, e, s_0)\right] = c_e. \tag{3}$$

The expected value integrates over the distribution $G\left(e\right)$ of permanent types and the distribution $F\left(s_0\right)$ of initial conditions. The timing in (3) allows some entrants to exit before production. Selection on the entry margin allows the composition of entering cohorts to change in response to prices. In particular, sufficiently low productivity entrants may exit immediately after drawing their productivity components—before producing and incurring any operating costs. The entry cost c_e is paid by all entrants, regardless of whether they choose to continue. We use the terms producing and non-producing entrants when we distinguish entrants that choose to operate and continue from entrants that choose to exit immediately.

[12] The ratio p_t/p_{t+1} converts the discount rate in goods to the numeraire.

[13] Sometimes this condition is formulated as a weak inequality, $E[V_t(0, e, s_0)] \leq c_e$. We will only consider parametrization with entry in each period, so this condition always holds with equality.

Measuring firm heterogeneity To keep track of heterogeneous firms, we let $\mu_t(a, E, S)$ be the measure in t of producing firms with age a and productivity components $e \in E$ and $s \in S$. These are the incumbents, $a \geq 1$, and the producing entrants, $a = 0$. We let $M_t > 0$ be the measure of all entrants (producing and non-producing). With market size L_t growing at rate η, we introduce measures $\bar{\mu}_t = \frac{\mu_t}{L_t}$ and $\bar{M}_t = \frac{M_t}{L_t}$ that are normalized by market size L_t.

Evolution of firm heterogeneity Let $\bar{\mu}_t(a, E, S)$ be the normalized measure in t of firms with age a and productivity components $e \in E$ and $s \in S$. The law of motion for $\bar{\mu}_t$ is:

$$\bar{\mu}_t(a, E, S) = \begin{cases} \iint_{e \in E, s \in S} (1 - x_t(0, e, s))\, G(de) F(ds) \bar{M}_t & a = 0 \\ \iiint_{e \in E, s \in S, s_{-1}} (1 - \delta)(1 - x_t(a, e, s))\, P(ds | s_{-1}) \frac{\bar{\mu}_{t-1}(a-1, de, ds_{-1})}{1+\eta} & a \geq 1 \end{cases}, \tag{4}$$

where s_{-1} is the stochastic productivity component last period. The piecewise formulation in age, a, accounts in turn for selection from total $\bar{M}$ into producing entrants and for selection among incumbents.[14]

Product market clearing Using the CES sector demand, and the normalized measure of producing firms, $\bar{\mu}_t$, market clearing requires:

$$p_t^{-\epsilon} = \sum_{a \geq 0} \iint_{e,s} y(es) \bar{\mu}_t(a, de, ds). \tag{5}$$

Stationary industry equilibrium We define the familiar stationary equilibrium featuring entry and exit with constant market size growth rate η. This is a partial equilibrium because the real interest rate, r, is taken as given. A *stationary industry equilibrium* is a constant price p, value function V with policy functions n and x, measure $\bar{\mu} = \mu_t / L_t$ over all firms, and measure $\bar{M} = M_t / L_t$ of entrants, both normalized by market size, such that (i) V solves Bellman equation (2), (ii) free entry condition (3) is satisfied, (iii) $\bar{\mu}$ and $\bar{M}$ solve the law of motion (4), and (iv) the product markets clears (5).

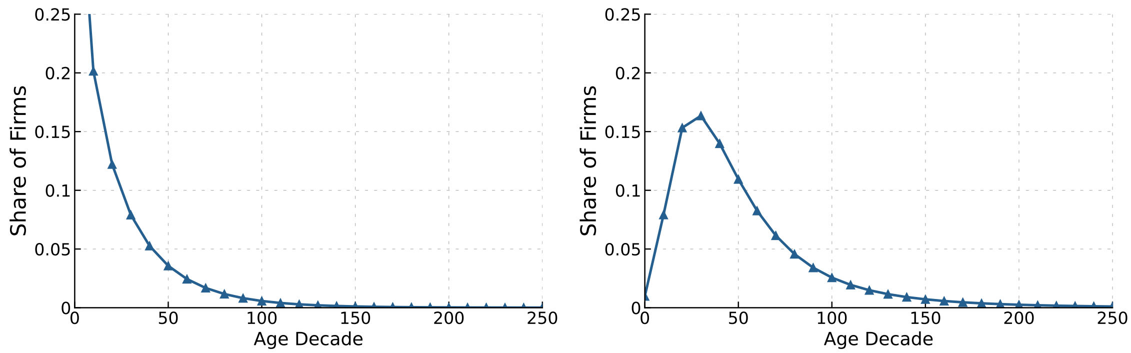
4.2 Firm Age Distribution

We plot the stationary age distribution in the baseline model in Figure 7, using parameter values from Karahan, Pugsley, and Şahin (2024) as an example. Panel A shows the age decade distribution among all firms, and Panel B shows the age decade distribution among the largest firms by revenue. We see that a hump-shaped age distribution can emerge among top firms even

[14] When solving the model, we assume there is some absorbing "maximum" age, $\bar{a}$, sufficiently large that the influence of initial condition s_{i0} is effectively zero. See Internet Appendix IA2.

Figure 7: Stationary Firm Age Distribution in Baseline Model: Overall and Top Firms

This figure plots the firm age distribution of the stationary equilibrium in the baseline model. We use the parametrization from Karahan, Pugsley, and Şahin (2024) with exogenous exit $\delta = 0.005$ and market size growth rate $\eta = 0.01$. Panel A shows the age decade distribution among all firms. Panel B shows the age decade distribution among the top 0.5% firms by revenue.



Panel A. All Firms

Panel B. Top Firms

in the stationary benchmark model. Through the lens of the productivity process in Equation (1), both high permanent, $\log e_i$, and persistent, $\sum_{k=0}^{a-1} \rho^k \sigma_\varepsilon \varepsilon_{it-k}$, components are necessary to become a top firm. As the persistent component takes some years to accumulate, the age distribution among top firms shows a hump-shaped pattern. Accordingly, we cannot use the clustering of birth years among top firms in a single cross section of data to detect special cohorts. Rather, we need the distribution of birth years over time. If special cohorts exist, the clustering of birth years will remain similar in *calendar year* for top firms at different points in time, and the *age distribution* of top firms will be nonstationary.

As we see in Section 3, top industrials feature a special generation. The stationary equilibrium of the baseline model, by definition, does not match these nonstationary empirical patterns. Recent work has considered deviation from the stationary equilibrium due to shifting demographics: slowing labor supply growth can reduce the entry rate, and in turn the share of young firms overall and among top firms (Karahan, Pugsley, and Şahin, 2024; Hopenhayn, Neira, and Singhania, 2022). However, such forces would affect all sectors, not just industrials; and the labor supply growth slowdown starting in the 1980s is too recent to explain the special generation among top industrials, which we already observe by the 1950s. In Section 5, we study how modifications of the baseline model can give rise to a special generation among top firms. In Section 6, we study what distinguishes industrials from retail/wholesale trade, where the firm age distribution appears stationary.

5 Industrials: Firm Dynamics with Special Generation

We examine three hypotheses for the origin of the special generation among top industrials: new industrial technology, time-varying market size growth, and time-varying entry costs. We extend the baseline model to implement each hypothesis as a perfect foresight shock that moves the industrial sector away from its stationary equilibrium. For each, we calibrate the model and the shock using the same set of moments, and analyze whether a given mechanism can sustain a special generation along the equilibrium transition path in line with the empirical facts.

5.1 New Industrial Technology

Business history research highlights two key features of the Second Industrial Revolution, which happened around the time when the special generation among top industrials emerged: the availability of modern capital intensive technology with higher returns to scale, and first-mover advantages due to improved efficiency with experience (Chandler, 1992, 1994). We extend the baseline model in Section 4 to incorporate both.

First, the modern technology uses capital to achieve larger scale production. This technology requires an adoption cost, which we assume was initially too high but declined during the Second Industrial Revolution, commonly dated from 1860 to 1900 (Gordon, 2000). *Entrants* can utilize this technology if they pay the adoption cost, whereas organizational or technological frictions make it difficult for incumbents to change technology (Arrow, 1974; Christensen, 1997; Carroll and Hannan, 2000; Jovanovic, 2001; Jovanovic and Rousseau, 2005; O'Reilly and Tushman, 2021; Hornbeck et al., 2024). According to historians, the large scale production required "new and improved processes of production developed for the first time in history," which were successfully implemented by entrants in U.S. and Germany but less so in Britain (Lamoreaux, 1988; Chandler, 1994; Mokyr, 2000).

Second, we allow firms to become more efficient with experience, such as by accumulating organizational capital, knowhow, or customer base (Ohyama, Braguinsky, and Murphy, 2004; Levitt, List, and Syverson, 2013). We refer to this process as "learning," but the mechanism can be interpreted broadly.[15] As Chandler (1994) explains: "The first movers were apt to be well

[15] Chandler (1992) elaborates that firms' "learned capabilities" result from "solving problems of scaling up the processes of production, from acquiring knowledge of customers' needs and altering product and process to services needs, coming to know the availabilities of supplies and the reliability of suppliers, and in becoming knowledgeable in the ways of recruiting and training workers and managers."

down the learning curve in each of the industry's functional activities before challengers went into full operation." Such learning stays in a given firm and cannot be easily replicated by others (i.e., it helps the firm relative to its competitors and does not have spillovers).

We examine whether a shock that reduces the adoption cost of the modern technology can generate a special generation among top industrials. With learning enhancing first-mover advantages for early adopters, we find that a long-lasting special generation can indeed emerge.

5.1.1 Setup

The setup modifies the baseline model in Section 4 in two ways: (i) we add a capital intensive modern technology that can be adopted at entry, and (ii) we let all firms learn from experience. We describe how we implement these features in the model. Then, we examine the response to a perfect foresight shock to the adoption cost of the modern technology that mimics the Second Industrial Revolution.

Capital intensive modern technology The modern technology, $y_{iat} = z_{iat} k_{iat}^{\alpha} n_{iat}^{\theta}$, incorporates capital with elasticity $\alpha > 0$. We assume $\alpha + \theta < 1$ to maintain diminishing returns to scale. In addition, we assume perfect rental markets for capital, with a constant rental rate R expressed in goods. Entrants choose whether to adopt the modern technology, with an adoption cost c_{Mt} that can vary over time.[16]

The capital intensity of the modern technology effectively increases the returns to scale in labor. To see this, we can write an expression for the modern technology that maximizes out over capital to characterize production net of capital costs as a function of only labor: $\max_k \left\{ z_{iat} k_{iat}^{\alpha} n_{iat}^{\theta} - Rk \right\} = \left(\chi z_{iat} n_{iat}^{\theta} \right)^{\frac{1}{1-\alpha}}$, with $\chi \equiv \frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{R^{\alpha}}$. The capital intensive technology has two effects on scale. If the rental rate is not too high, it boosts the productivity of labor over the traditional technology, i.e., $\chi > 1$. Most importantly, it increases the returns to scale in labor. In effect, the span of control parameter increases from θ to $\frac{\theta}{1-\alpha}$. For large enough z_{it}, even if $\chi < 1$, a firm operating the modern technology will be significantly larger. The technology heterogeneity captures features of the growth rate heterogeneity studied in Luttmer (2011) and Guntin and Kochen (2025).

[16] We stipulate that firms choose the technology at entry and stay with the decision afterwards (e.g., they cannot change the technology later, either internally or by acquiring other firms) because the technology also requires accompanying organizational processes that are difficult to change once a company has been established (Lamoreaux, 1988; Chandler, 1994). New organizational structures that accompany a different technology are easier to adopt by new firms.

Learning from experience All firms, regardless of technology, become more efficient in production with time. Now, $z_{iat} = e_i \exp(\zeta a) s_{iat}$. This introduces a learning component to productivity that grows deterministically at rate ζ until maximum age $\bar{a}$. With learning, log productivity for firm i of age a in period t written as a moving average is:

$$\log z_{iat} = \log e_i + \zeta a + \rho^a \log s_{i0} + \sum_{k=0}^{a-1} \rho^k \sigma_\varepsilon \varepsilon_{it-k}. \tag{6}$$

Model implementation Incorporating these features requires allowing for both firms using modern technology (“modern firms") and firms using traditional technology (“traditional firms"), specifying the technology adoption choice, and properly aggregating over the types.

First, we characterize the value of a firm using the modern technology recursively, and with it the optimal labor and exit choices. Let $V_t^M(a, e_i, s_{iat})$ be the value in period t of modern firm i with age a, permanent productivity component e_i, and stochastic component s_{iat}. Then V_t^M satisfies the following Bellman equation:

$$V_t^M = \max_{x \in \{0,1\}} \left\{ \max_n \left\{ p_t \left(\chi z_{iat} n^\theta \right)^{\frac{1}{1-\alpha}} - n \right\} - c_f + \frac{1-\delta}{1+r} \frac{p_t}{p_{t+1}} E_t \left[V_{t+1}^M \right], 0 \right\}. \tag{7}$$

This Bellman equation for V_t^M is the modern technology counterpart to equation (2) for V_t with the traditional technology. Optimal labor, n_t^M, and exit, x_t^M, are also defined analogously.

Second, we compare the value of entering firms with the modern versus the traditional technology to determine the technology adoption decision. Potential entrants now evaluate the expected value of entry incorporating the possibility of adopting the modern technology. Conditional on drawing a permanent e_i and initial s_{i0}, a firm can use c_{Mt} units of labor to implement the modern technology or use the traditional technology at no additional cost. Incorporating the adoption decision, the free entry condition becomes:

$$E \left[\max_{m \in \{0,1\}} \left\{ V_t(0, e, s_0), V_t^M(0, e, s_0) - c_{Mt} \right\} \right] = c_e. \tag{8}$$

Note that the adoption decision is inside the expectation, because it is conditional on the entering firm's draw from distribution $G(e)$ over permanent types and $F(s_0)$ over initial conditions. Policy function $m_t(e, s_0)$ captures the adoption decision: it equals one if an entering firm with productivity components e and s_0 chooses to adopt the modern technology and zero otherwise.

Third, we need a new measure to track modern firms and aggregate over both modern and traditional firms. Let $\bar{\mu}_t^M(a, E, S)$ be the measure of age a modern firms with productivity $e \in E$ and $s \in S$, also normalized by market size. In light of the technology adoption decision,

m_t, we adjust the law of motion in Equation (4) for traditional firms as follows:

$$\bar{\mu}_t(a,E,S) = \begin{cases} \iint_{e\in E, s\in S} (1 - x_t(0,e,s))(1 - m_t(e,s)) G(de) F(ds) \bar{M}_t & a = 0 \\ \iiint_{e\in E, s\in S, s_{-1}} (1-\delta)(1 - x_t(a,e,s)) P(ds|s_{-1}) \frac{\bar{\mu}_{t-1}(a-1,de,ds_{-1})}{1+\eta} & a \geq 1 \end{cases} \tag{9}$$

This accounts for the technology adoption choice at entry, i.e., when $a = 0$. We then define analogously a new law of motion for the measure of modern firms, $\mu_t^M(a,E,S)$:

$$\bar{\mu}_t^M(a,E,S) = \begin{cases} \iint_{e\in E, s\in S} \left(1 - x_t^M(0,e,s)\right) m_t(e,s) G(de) F(ds) \bar{M}_t & a = 0 \\ \iiint_{e\in E, s\in S, s_{-1}} (1-\delta)\left(1 - x_t^M(a,e,s)\right) P(ds|s_{-1}) \frac{\bar{\mu}_{t-1}^M(a-1,de,ds_{-1})}{1+\eta} & a \geq 1 \end{cases} \tag{10}$$

In both laws of motion, the symbol $\bar{M}$ still denotes the normalized measure of all entrants (producing and non-producing); it does not refer to the modern technology.

Finally, we close the model by adjusting the market clearing condition to include output of both traditional and modern firms. Market clearing becomes:

$$p_t^{-\epsilon} = \sum_{a\geq 0} \iint_{e,s} \left(y_t(e\exp(\zeta a)s)\bar{\mu}_t(a,e,s) + y_t^M(e\exp(\zeta a)s)\bar{\mu}_t^M(a,e,s)\right) de\, ds. \tag{11}$$

Transition path following modern technology shock We start from a stationary equilibrium in 1859 with no incumbents using the modern technology due to high adoption costs.[17] The *modern technology shock* is a perfect foresight path for c_{Mt} since 1860 that declines exponentially for 40 years from c_{M1860} to c_{M1899}, and remains constant thereafter. This timing follows the standard dating of the Second Industrial Revolution (Gordon, 2000). The equilibrium following the shock is a sequence of prices $\{p_t\}_{t=1860}^{\infty}$, value functions $\{V_t, V_t^M\}_{t=1860}^{\infty}$ with corresponding policy functions $\{n_t, x_t, n_t^M, x_t^M\}_{t=1860}^{\infty}$, market-size scaled measures over all firms $\{\bar{\mu}_t, \bar{\mu}_t^M\}_{t=1860}^{\infty}$, and potential entrants with technology adoption policy $\{\bar{M}_t, m_t\}_{t=1860}^{\infty}$, such that for $t \geq 1860$: (i) V_t and V_t^M solve Bellman equations (2) and (7), respectively, (ii) free entry condition (8) incorporating the technology adoption choice is satisfied, (iii) $\bar{\mu}_t$, $\bar{\mu}_t^M$, and $\bar{M}_t$ satisfy the two laws of motion (9) and (10), and (iv) the product market clears (11).

5.1.2 Calibration

To calibrate this model with heterogeneous technologies and learning, we need to set 19 parameters. There are 15 parameters that govern technology, discounting, costs, and productivity

[17] The stationary equilibrium is defined analogously to Section 4 except it allows for both traditional and modern firms' value functions, policy functions, and measures. It uses the laws of motion (9) and (10), free entry condition (8), and market clearing (11), which all reflect the choice over traditional and modern technology.

Table 1: Externally Calibrated Parameters

This table shows the externally calibrated parameters, which are held constant in all calibrations (except the analysis in Section 5.2 calibrates the market growth rate η as a shock process).

Definition	**Parameter**	**Value**	**Definition**	**Parameter**	**Value**
Real interest rate	r	0.04	Exogenous exit rate	δ	0.004
Rental rate	R	0.1	Elasticity of demand	ϵ	0
Span of control	θ	0.64	Market growth rate	η	0.01

dynamics; two parameters control demand for industrial sector output; and two parameters set the starting and finishing levels of adoption costs for the modern technology shock. We fix six parameters externally and calibrate the remaining 13 by matching a set of data moments to the model along its *transition* path. We explain our calibration strategy in detail below.

Table 1 reports the values of the six externally calibrated parameters. The period length is one year, and in our industry equilibrium setting, we set the real interest rate $r = 0.04$ and the rental rate $R = 0.10$, which assumes a six percent annual depreciation rate. We set the span of control $\theta = 0.64$ to match the average labor share, and we set an exogenous exit rate $\delta = 0.004$ to match the annual exit rate of industrial firms with 500 or more employees measured in the BDS over 2014 to 2018. Finally, we let sector demand be inelastic, i.e., $\epsilon = 0$, with market size L_t growing at a constant rate of one percent, i.e., $\eta = 0.01$.

Table 2, Panel A, columns (1) and (2) list the 13 internally calibrated parameters, which are chosen by matching a set of moments along the model's transition path to their data counterparts following a "one-step" procedure. That is, we calibrate the parameters governing the stationary distribution and the modern technology shock simultaneously: conditional on a full set of parameters, we first solve for an initial stationary equilibrium where the adoption cost is sufficiently high so no firm adopts the modern technology, and then for the transition path in response to a parameterized perfect foresight modern technology shock that reduces the modern technology adoption cost. This one-step procedure is important because, as we emphasize in Section 3, we never observe the industrial sector during a fully stationary period. Although some data moments, such as those conditional on age, may be stable (Karahan, Pugsley, and Şahin, 2024), others are not: the median age of top firms, for example, changes slowly over time due to the special generation.

Table 2, Panel B, column (1) lists the set of moments. We choose these moments to jointly discipline features of the stationary equilibrium and the transition path. To start, we have nine

Table 2: Calibration for Industrials

Panel A shows the calibrated parameters for the industrial sector in the three different models in Sections 5.1, 5.2, and 5.3. The initial stochastic component of productivity, s_{i0}, is drawn from a distribution $F(s) \sim \text{LogNormal}(\mu_F, \sigma_F^2)$. The permanent component of productivity, e_i, is drawn from an adjusted lognormal distribution $G(e)$ that includes a superstar state to increase fat-tailedness (Internet Appendix IA2.1). In column (3), the modern technology shock is calibrated so the adoption cost of the modern technology declines exponentially for 40 years starting in 1860, from an initial adoption cost of c_{M1860} to c_{M1899}, and remains constant thereafter. In column (4), the market size grows at rate η_{1859} in the initial steady state, at η_{1860} from 1860 to 1969, and at η_{1970} thereafter. In column (5), the initial steady-state entry cost is c_{e1859}. From 1859 to $1859 + \tau_{c_e}$, the entry cost follows an increasing sigmoid function in equation (13) from c_{e1859} to $c_{e1859+\tau_{c_e}}$, and remains constant thereafter. Panel B shows a comparison of data and model moments across these three models. Growth rates, exit rates, and top firm revenue shares are averaged over the 2014 to 2018 period using Census Revenue LBD data. Special generation share is the share of top firms born between 1880 and 1919. Top firms are the top 0.15% by revenue. Revenue LBD data moments are calculated on FSRDC Project Number 1731. (CBDRB-FY25-P1731-R12291)

Panel A. Internally Calibrated Parameters

Definition (1)	**Parameter** (2)	**Modern Technology** (3)	**Market Growth** (4)	**Entry Cost** (5)
Entry cost	c_e	24.370	42.222	—
Overhead cost	c_f	3.575	9.004	58.477
Persistence of AR(1)	ρ	0.915	0.955	0.962
Volatility of AR(1)	σ_ε	0.175	0.189	0.450
Initial condition mean	μ_F	-0.113	-0.778	-0.040
Initial condition dispersion	σ_F	0.388	0.709	1.371
Permanent type dispersion	σ_G	0.135	0.104	0.070
Superstar probability	p_G^*	0.001	0.001	0.005
Superstar factor	λ_G^*	2.981	8.108	9.373
Learning rate	ζ	0.003	0.004	0.002
Modern capital elasticity	α	0.144	—	—
Initial adoption cost	c_{M1860}	27,000	—	—
Terminal adoption cost	c_{M1899}	8,000	—	—
Initial market growth rate	η_{1859}	—	0.0003	—
High market growth rate	η_{1860}	—	0.018	—
Low market growth rate	η_{1970}	—	0.010	—
Initial entry cost	c_{e1859}	—	—	5.002
Terminal entry cost	c_{e2061}	—	—	26.609
Entry cost steepness	k	—	—	5.451
Entry cost length	τ_{c_e}	—	—	202

Panel B. Calibration Moments

Moment (1)	**Data** (2)	**Modern Technology** (3)	**Market Growth** (4)	**Entry Cost** (5)
Age 1–5 revenue growth rate	0.041	0.023	0.034	0.042
Age 6–10 revenue growth rate	0.004	0.012	0.027	0.001
Age 11–15 revenue growth rate	0.018	0.008	0.021	0.008
Age 1–2 exit rate	0.136	0.121	0.120	0.183
Age 3–5 exit rate	0.098	0.065	0.058	0.097
Age 6–10 exit rate	0.074	0.046	0.047	0.056
Age 11–15 exit rate	0.053	0.030	0.033	0.034
Top 1% entrants revenue share	0.30	0.36	0.38	0.46
Top firms revenue share	0.63	0.55	0.51	0.49
Top firms median age in 2018	98	100	96	101
Top firms median age in 1955	61	59	63	57
Special generation share in 2018	0.35	0.38	0.27	0.31
Special generation share in 1955	0.55	0.56	0.41	0.47

moments that capture the overall firm dynamics around 2018, building on Karahan, Pugsley, and Şahin (2024): exit rates by age group, revenue growth rates by age group, and the revenue share of the top 1% entrants in a year. These moments, which are all conditional on firm age, have been relatively stable over recent decades. Karahan, Pugsley, and Şahin (2024) used these moments conditional on age rather than other moments like the overall exit rate or entry rate that can change over time with the firm age composition. We measure these moments for the industrial sector using the Census LBD Revenue dataset, and take the average between 2014 and 2018 to smooth out short-term fluctuations. We use these moments primarily to inform the parameters that govern the "classical" components of the model: the entry cost c_e, the overhead cost c_f, the persistence of the productivity shock ρ, the volatility of the productivity shock σ_ε, as well as the mean, μ_F and standard deviation, σ_F for the log normal distribution $F(s_0)$ of its initial condition. These classical moments also inform three additional parameters for the distribution over permanent types $G(e)$.

The distribution $G(e)$ over permanent type is constructed to allow for fat-tailedness. We start by approximating a mean zero lognormal distribution with standard deviation σ_G using three symmetric points in logs. Then, we incorporate a fourth superstar grid point that is λ_G^* times larger than the largest (3rd) grid point and is drawn with probability p_G^*. In other words, with probability $1 - p_G^*$, type e is drawn from the approximately log normal distribution. We refer to p_G^* as "superstar probability" and λ_G^* as "superstar factor" going forward (see Internet Appendix IA2.1 for further details). The adjustment allows for a fat tail and helps match the observed firm size distribution, which is highly skewed. In combination with heterogeneous technologies, the fat tailed distribution of permanent types allows for the patterns of rapid growth in Luttmer (2011) and Guntin and Kochen (2025).

Next, we include five moments about top firms: the median age of top firms in 2018 and 1955, the fraction of top firms born between 1880 and 1919 (the special generation share), and top firms' share of all revenue in 2018. These moments along the transition path are informative about the modern technology (governed by capital elasticity α), the path of adoption costs c_{M1860} and c_{M1899}, and the learning rate ζ. In the data, top firms are the top 388 industrials by revenue because there are 388 of them in the 2018 *Fortune* 1,000 companies. We approximate them in the model using the top 0.15% firms by revenue because there are around 250,000 industrial firms in 2018 (according to the BDS dataset) and in 1955 (according the IRS Statistics of Income data). We measure the median age and special generation shares in 2018 and 1955 using our dataset from Section 2. For the top firm revenue share, we rely on the Census Bureau's Revenue LBD database (averaged over 2014 to 2018).

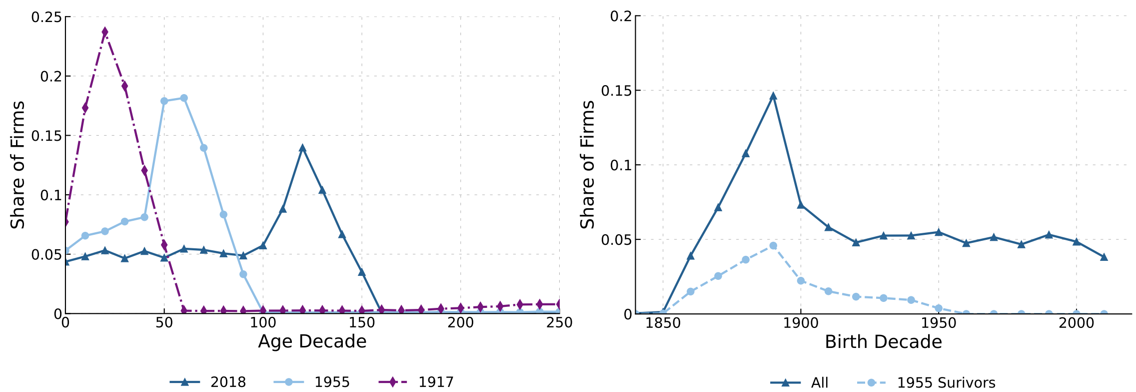
Table 2 shows the values of internally calibrated parameters in Panel A column (3), and model moments in Panel B column (3). First, the modern technology adoption cost starts high in 1860 at 27,000 units of labor, and declines considerably over the next four decades to 8,000 units of labor, which is still substantial (median firm size is about nine units of labor); its path is shown in Figure IA9. About 0.1% of entrants obtain the superstar permanent type, and they adopt the modern technology. Figure IA11 shows that the modern technology's adoption cost in the stationary equilibrium in 1859 is sufficiently high that no firm is using it, and then entrants using the modern technology gradually emerge in the following decades. Second, the degree of learning is modest, yet it still helps the longevity of the special generation as we show below. Finally, among the more classical parameters, the persistence ρ and volatility σ_ε of the stochastic component of productivity are close to but slightly lower than common values in the literature (Sterk, Sedláček, and Pugsley, 2021). The dispersion of the permanent type σ_G and the initial condition σ_F are generally high in our calibrations to match the top entrant revenue share. The initial condition tends to have a low mean μ_F to generate enough exit.

5.1.3 Results

Figure 8 compares the key empirical facts in the data with their counterparts in the calibrated model. Panel A plots the age distribution of the largest firms in 2018, 1955, and 1917, which is the model counterpart of Figure 3 for U.S. industrials in the data. We find that the availability of the modern technology leaves a strong and concentrated imprint on the age distribution that fades only very slowly with time. In 1917, the largest firms are very young. These are the early adopters of the modern technology as its cost declines. This early cohort remains the most important among the top firms in 1955 and 2018. Panel B plots the birth decade of the top firms in 2018 (solid line) and the subset of firms that are also the top firms in 1955 (dashed line), which is the model counterpart to Figure 4. Like in the data, the largest firms in 2018 are disproportionately drawn from the cohort born around 1900, and a modest fraction of the largest firms in 2018 are also among the largest firms in 1955. Given the volatility of the productivity shock (σ_ε) required to match the exit rate and growth rate dynamics, individual firms at the top experience a reasonable amount of churn. Meanwhile, the capital intensive technology gives them some staying power. These two forces balance to produce a degree of overlap between top firms in 1955 and 2018 similar to the data.

Figure 8: Top Industrials in Model with Modern Technology Shock and Learning

Panel A is the model-implied age decade distribution of the top 0.15% industrial firms in 2018, 1955, and 1917. Panel B is the model-implied birth decade distribution of the top 0.15% industrial firms in 2018 and the subset that are also among the top 0.15% industrial firms in 1955. The model uses parameters in column (3) of Table 2, Panel A. We simulate the model 500 times, each with an initial 100K firms drawn from the ergodic distribution, a 500-period burn-in, and then for 160 periods following the shock. Results are averaged over 500 simulations.



Panel A. Age Distribution of Top Firms

Panel B. Top Firms in 2018 and 1955

Overall, this model produces a special generation that matches the empirical facts, including the age distributions among top firms and the overlap among top firms over time, which we do not explicitly target. Matching the evidence relies on three key elements. First, the special generation needs to have advantages relative to firms that came before. This arises from their adoption of the modern technology, which amplifies their scale. Second, the special generation needs to have lasting advantages relative to potential entrants afterwards. This arises from time-varying adoption costs combined with learning that enhances the first-mover advantage, which we elaborate below. Third, to achieve a persistent cohort together with churning among individual companies at the top, we need the right mix of ex ante high permanent type and ex post accumulated good luck through the stochastic component. The high permanent type incentivizes firms to adopt the modern technology; it also makes potential su perstars from the special generation, even those that are not dominant early, more likely to survive long enough to accumulate a sufficiently large ex post component and become dominant later. The stochastic component of productivity modulates churning of individual companies.

We further unpack the second component regarding the lasting advantage of the special generation, which depends on the relatively swift decline of the modern technology adoption cost as well as learning. The 40 year decline of the adoption cost mimics the timing of the Second Industrial Revolution. It creates at first a surge in entering firms while the cost declines, but then a sustained reduction in entering firms after the adoption cost reaches its permanently lower level, as shown in Figure IA10.[18] This gives rise to a special generation clustered in time. In Figure IA13, we stretch out the adoption cost decline to over 100 years, keeping other parameters including the beginning and ending level of the adoption cost the same as before. In this case, adopters of the modern technology emerge too gradually and many of them too late to produce a clustering of top firms born around 1900. These variations could shed light on why special generations did not appear among top British industrials: according to Chandler (1994), British companies were slow to embrace modern production and management structures.[19]

Learning from experience also enhances the first-mover advantage of the special generation, as the early adopters of the modern technology continue to accumulate experience when entry remains low. Without learning, highly productive late modern entrants would be no different than similarly productive early entrants. It turns out that even a modest rate of learning, $\zeta = 0.003$ (i.e., productivity increases by around 0.3% with every additional year of experience), can considerably extend the longevity of the special generation. In Figure IA14, we examine the same shock without learning, $\zeta = 0$ (using the same set of parameters as in Table 2, Panel A, column (3) except for ζ). In this case, the echo of the early cohort is still detectable in 2018, but the cohort is no longer the most represented, because its importance fades more quickly. The head start from gradual learning gives the early entrants a much longer-lasting advantage.

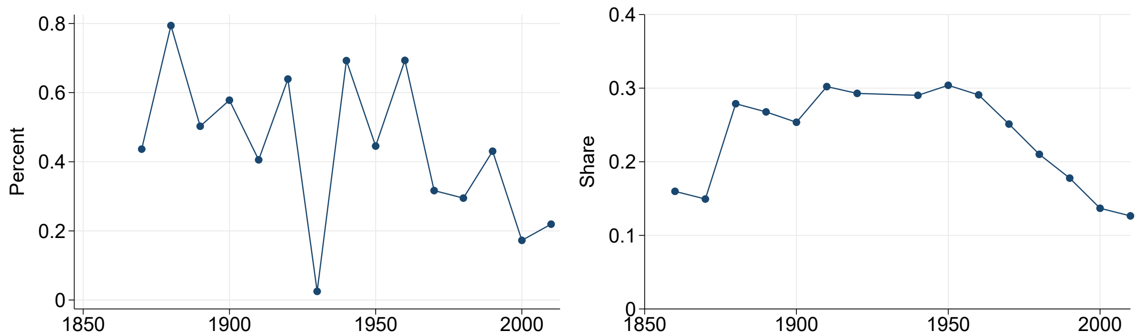
The modern technology shock that fits the patterns in the data generates top firm dynamics that remain nonstationary for quite some time. While the model is stationary in the very long run, Figure IA12 shows the special generation among top industrials is still present 250 years after the shock. Only after 500 years does the age distribution of top industrials approach stationarity. These results suggest that top firm dynamics may not be well approximated by a stationary equilibrium, and technological change can create special cohorts that persist for centuries even with competitive markets.

[18] During the first phase, the downward pressure on prices from the availability of the modern technology (Panel A of Figure IA10) creates a surge in entry, as low productivity incumbents exit and surviving incumbents shrink (Panel B of Figure IA10). Over this period, both the propensity to adopt the modern technology and the size of the entering cohort increase. During the second phase, as adoption costs reach their minimum and prices are no longer declining, fewer entrants are required to clear the product market. Even as the propensity to adopt the modern technology stays at its highest, the size of the entering cohorts remain low.

[19] Chandler (1994) writes: "What differentiated British entrepreneurial, later family-controlled, enterprises from those in the United States and Germany was that the entrepreneurs assembled smaller management teams, and until well after World War II they and their heirs continued to play a larger role in the making of middle and top-management decisions." The insistence on personal and family management could have contributed to higher effective costs of operating modern large-scale industrial enterprises.

Figure 9: Manufacturing Growth

This figure shows value added growth (Panel A) and value added shares (Panel B) of the manufacturing sector by decade. Historical data are from Carter et al. (2006) and historical national accounts provided by the Groningen Growth and Development Centre. Recent data are from the BEA.



Panel A. Real Value Added Growth

Panel B. Value Added Share

5.2 Time-Varying Market Size Growth

One intuition is that the disproportionate importance of a special generation among top firms can arise from rapid growth in market size. Since fluctuations in market size under free entry are primarily accommodated through entry (Hopenhayn, Neira, and Singhania, 2022; Karahan, Pugsley, and Şahin, 2024), a surge in the growth rate of market size could induce unusually large cohorts of entrants. In the presence of learning that enhances staying power, the larger cohort may result in a special generation. In this account, the productivity distribution over potential entrants remains the same, and additional superstars emerge through a period with more "shots on goal." Figure 9 shows that U.S. manufacturing experienced high growth until 1970, and low growth afterwards. In this section, we examine whether a shock to the market size growth rate of this nature can explain the empirical facts.

5.2.1 Setup

The setup is identical to the baseline model in Section 4, except for two modifications: (i) the market size growth rate η_t varies over time, and (ii) we allow for learning as in Section 5.1.

Time-varying market size growth rate We allow the growth rate η of market size L_t in the CES demand function to vary over time. That is, we now have $L_t = (1 + \eta_t)L_{t-1}$.

Learning from experience As in Section 5.1, firms become more efficient in production with time at rate ζ (until maximum age $\bar{a}$). The evolution of log productivity for firm i of age a

in period t follows the moving average given by Equation (6).

Model implementation The value function is the same as in the baseline model and satisfies Bellman equation (2). Product market clearing is analogous to Equation (5) in the baseline model, except we need to account for productivity increasing at rate ζ due to learning:

$$p_t^{-\epsilon} = \sum_{a \geq 0} \iint_{e,s} y_t(e \exp(\zeta a) s) \bar{\mu}_t(a, e, s) de\, ds. \tag{12}$$

The evolution of the firm distribution is identical to Equation (4) in the baseline model, except for the time-varying growth rate η_t in market size.

Transition path following market size growth shock Starting from a stationary equilibrium in 1859 (as in Section 4) with η_{1859}, the *market size growth shock* is a perfect foresight sequence of market size growth rates $\{\eta_t\}_{t=1860}^{\infty}$ that are elevated at $\eta_{1860} > \eta_{1859}$ for 110 years and then decline to η_{1970}, remaining constant thereafter. The equilibrium following the shock $\{\eta_t\}_{t=1860}^{\infty}$ is a sequence of prices $\{p_t\}_{t=1860}^{\infty}$, value functions $\{V_t\}_{t=1860}^{\infty}$ with corresponding policy functions $\{n_t, x_t\}_{t=1860}^{\infty}$, market-size scaled measures over all firms $\{\bar{\mu}_t\}_{t=1860}^{\infty}$, and entrants $\{\bar{M}_t\}_{t=1860}^{\infty}$, such that for $t \geq 1860$: (i) V_t solves Bellman equation (2), (ii) free entry condition (3), (iii) $\bar{\mu}_t$ and $\bar{M}_t$ satisfy the law of motion (4) with time-varying growth rate η_t in market size, and (iv) the product markets clears (12).

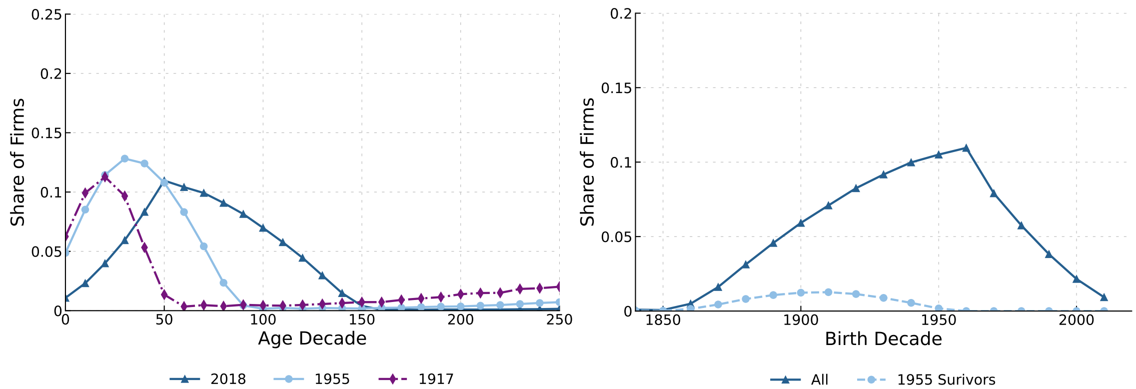
5.2.2 Calibration

We calibrate this model to the data following the same one-step strategy as in Section 5.1.2. We keep the values of the externally calibrated parameters constant. To ease comparisons across calibrations, we use the same set of data moments shown in column (1) of Table 2, Panel B. For the model parameters, the first ten are the same as before. The final three model parameters now govern the path of market size growth rather than the modern technology. They include the initial market size growth rate η_{1859}, the faster market size growth rate η_{1860} from 1860 to 1969, and the low market size growth rate η_{1970}.

Table 2 shows the values of internally calibrated parameters in Panel A column (4), and model moments in Panel B column (4). First, the market size growth rate starts at $\eta_{1859} = 0.0003$ and rises to the higher level of $\eta_{1860} = 0.018$, which persists for a century before falling to $\eta_{1970} = 0.010$; the full path is shown in Figure IA15. Second, about 0.1% of entrants obtain the superstar permanent type like before, while the superstar factor is higher than that in Section 5.1. Without technological heterogeneity, a longer tail of the permanent type is needed to match

Figure 10: Top Industrials in Model with Market Size Growth Shock and Learning

Panel A is the model-implied age decade distribution of the top 0.15% industrial firms in 2018, 1955, and 1917. Panel B is the model-implied birth decade distribution of the top 0.15% industrial firms in 2018 and the subset that are also among the top 0.15% industrial firms in 1955. The model uses parameters in column (4) of Table 2, Panel A. We simulate the model 500 times, each with an initial 100K firms drawn from the ergodic distribution, a 500-period burn-in, and then for 160 periods following the shock. Results are averaged over 500 simulations.



Panel A. Age Distribution of Top Firms

Panel B. Top Firms in 2018 and 1955

top firms' revenue shares among entrants and overall. Third, the degree of learning is similar to that in Section 5.1.

5.2.3 Results

Figure 10 compares the key empirical facts in the data with their counterparts in the calibrated model. Panel A plots the age distribution of the largest firms in 2018, 1955, and 1917, which is the model counterpart of Figure 3 for U.S. industrials. While top firms in the model exhibit aging over time, by 2018 the shape of the firm age distribution differs substantially from that in the data. In 2018, the modal top firm is around 50 years old (born around the 1960s) in this calibration versus closer to 100 (born around 1900) in the data. In the model, the largest cohort among top firms in 2018 comes from the last years of the higher market size growth era. The later cohorts are smaller as slower market size growth requires lower entry rates, while the older cohorts are less prevalent due to more attrition.[20] Panel B plots the birth decade of the top firms in 2018 (solid line) and the subset of firms that are also the top firms in 1955 (dashed

[20] Panel A of Figure IA16 shows that prices remain unaffected by the market size growth shock. This is because the free entry condition (3) that determines prices is independent of market size growth (Hopenhayn, Neira, and Singhania, 2022; Karahan, Pugsley, and Şahin, 2024). Instead, adjustment to the faster market size growth occurs through only firm entry (see Panel B of Figure IA16). Entry spikes initially and then stabilizes at an elevated level. After the high growth era ends, entry drops sharply before settling at its new steady state level.

line), which is the model counterpart to Figure 4. With learning, the model is able to generate a reasonable degree of overlap between top firms in 1955 and 2018, but the overall birth decade distribution, as we discuss above, is inconsistent with the data.

Overall, the prolonged high entry makes the cohort from the 1960s the dominant one in subsequent years. In other words, the observed path of manufacturing growth makes it challenging to generate a special generation of top firms clustered around 1900.

5.3 Time-Varying Entry Costs

Another intuition is that regulatory changes (e.g., weaker antitrust, costlier compliance) could lead to rising entry costs over time, which would diminish the size of later entering cohorts and entrench incumbents (Gutiérrez, Jones, and Philippon, 2021; Singla, 2023). In this section, we examine whether rising entry costs can produce the empirical facts.

5.3.1 Setup

The setup is identical to the baseline model in Section 4, except for two modifications: (i) we allow the entry cost c_e to increase over time, and (ii) we allow for learning as in Sections 5.1 and 5.2.

Time-varying entry cost To parametrize the path of the entry cost, c_{et}, we use an increasing sigmoid function:

$$c_{et} = \frac{A}{1 + e^{-k\tilde{t}}} + B. \tag{13}$$

The variable $\tilde{t} = \frac{2(t-1859)}{\tau_{c_e}} - 1$ maps calendar year t onto $[-1, 1]$. This allows the entry costs to increase smoothly starting at 1860 over a period of τ_{c_e} years and remain constant thereafter. A is the amplitude parameter and B is the level parameter. The sigmoid function creates a S-shaped path for entry costs, which is strictly increasing, continuous, differentiable, and non-negative.

Learning from experience As in Section 5.1, firms become more efficient in production with time at rate ζ (until maximum age $\bar{a}$). The evolution of log productivity for firm i of age a in period t follows the moving average given by Equation (6).

Model implementation The value function is the same as in the baseline model and satisfies Bellman equation (2). The free entry condition in Equation (3) is modified to allow for time variation in the entry cost, i.e., $E[V_t(0, e, s_0)] = c_{et}$. Product market clearing is the same as Equation (12) in Section 5.2. The evolution of the firm distribution is the same as Equation

(4) in the baseline model.

Transition path following entry cost shock Starting from a stationary equilibrium (as in Section 4) set in 1859 with constant entry cost c_{e1859}, the *entry cost shock* is a perfect foresight sequence of entry costs $\{c_{et}\}_{t=1860}^{\infty}$ that follow the increasing sigmoid function in Equation (13) for τ_{c_e} years and remains constant thereafter. The equilibrium following the shock $\{c_{et}\}_{t=1860}^{\infty}$ is a sequence of prices $\{p_t\}_{t=1860}^{\infty}$, value functions $\{V_t\}_{t=1860}^{\infty}$ with corresponding policy functions $\{n_t, x_t\}_{t=1860}^{\infty}$, market-size scaled measures over all firms $\{\bar{\mu}_t\}_{t=1860}^{\infty}$, and measures of entrants $\{\bar{M}_t\}_{t=1860}^{\infty}$, such that for $t \geq 1860$: (i) V_t solves the Bellman equation (2), (ii) the free entry condition (3) holds with the time-varying c_{et}, (iii) the product market clears as in (12), and (iv) $\bar{\mu}_t$ and $\bar{M}_t$ satisfy the law of motion (4).

5.3.2 Calibration

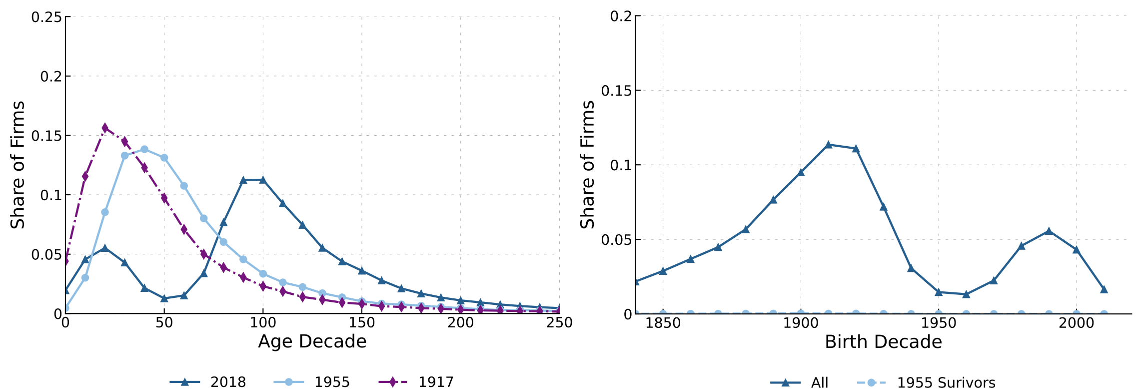
We calibrate this model to the data following the same one-step strategy as in Section 5.1.2. The values of the externally calibrated parameters remain the same. We continue to use the same set of data moments shown in column (1) of Table 2, Panel B. For the model parameters, the entry cost c_e becomes c_{e1859} and the three modern technology parameters (or three market size growth parameters) are replaced by $c_{e1859+\tau_{c_e}}$, k, and τ_{c_e}. These four parameters pin down the values of A and B in Equation (13).[21]

Table 2 shows the values of internally calibrated parameters in Panel A column (5), and model moments in Panel B column (5). First, for the path of the entry cost, Figure IA17 shows a gradual rise since 1860, accelerating around 1900, stabilizing around the 2000s, and ending in 2061, with the most rapid rise around the mid-1900s. The acceleration of rising entry costs needs to begin in the early 1900s to make decades around 1900 most represented among top firms subsequently. Second, the volatility of the stochastic component of productivity σ_ε is very high and so is the overhead cost c_f, which are needed to make top firms in 1955 sufficiently young. Otherwise, since rising entry costs per se do not phase out older incumbents, there can be too many top firms in 1955 from earlier in the 1800s. High volatility of productivity increases churning and high overhead costs generate more exit, which speed up the attrition of older incumbents. Finally, the learning rate of 0.002 is slightly lower than the level in previous calibrations. Higher learning rates would also contribute to top firms being too old in 1955.

[21] Solving equation (13) using boundary conditions c_{e1859} and $c_{e1859+\tau_{c_e}}$ for A and B yields: $A = \frac{c_{e1859+\tau_{c_e}} - c_{e1859}}{\frac{1}{1+e^{-k}} - \frac{1}{1+e^{k}}}$ $\quad B = c_{e1859} - \frac{A}{1+e^{k}}$.

Figure 11: Top Industrials in Model with Entry Cost Shock and Learning

Panel A is the model-implied age decade distribution of the top 0.15% industrial firms in 2018, 1955, and 1917. Panel B is the model-implied birth decade distribution of the top 0.15% industrial firms in 2018 and the subset that are also among the top 0.15% industrial firms in 1955. The model uses parameters in column (5) of Table 2, Panel A. We simulate the model 500 times, each with an initial 100K firms drawn from the ergodic distribution, a 500-period burn-in, and then for 160 periods following the shock. Results are averaged over 500 simulations.



Panel A. Age Distribution of Top Firms

Panel B. Top Firms in 2018 and 1955

5.3.3 Results

Figure 11 compares the key facts in the data with their counterparts in the calibrated model. Panel A plots the age distribution of the largest firms in 2018, 1955, and 1917, which is the model counterpart of Figure 3 for U.S. industrials in the data. With the rise of entry cost accelerating around the early 1900s, there can be an over-representation of top firms born before then.[22] Panel B plots the birth decade of the top firms in 2018 (solid line) and the subset of firms that are also the top firms in 1955 (dashed line), which is the model counterpart to Figure 4. In this case, the overlap between top firms across 2018 and 1955 is minimal. As mentioned above, the special generation needs to have advantage not only relative to potential entrants afterwards (due to rising entry costs), but also relative to firms that came before. The latter does not follow directly from rising entry costs, and needs to arise from high churn (e.g., high volatility of the stochastic component of productivity), which in turn suppresses the overlap between top firms in 1955 to 2018. This lack of overlap might appear to contradict the intuition that rising entry costs make incumbents more persistent. The core issue here is that rising entry costs alone

[22] Rising entry costs are accompanied by rising prices, shown in Panel A of Figure IA18. The startup rate falls as price growth accelerates (Panel B of Figure IA18), because higher prices diminish the number of entrants needed to clear markets. As the growth in prices slows, starting around 1950, the startup rate begins to recover. The startup rate could not continue to decline unless rising entry costs were to accelerate indefinitely.

would favor all incumbents, not specifically those born after the late 1800s, so the model needs a force to cleanse the older incumbents and that force produces too much churn.

Overall, we find that time-varying entry costs alone have difficulty matching the evidence. While common intuition might suggest that rising entry costs could contribute to the largest firms in 2018 being old, the special generation requires more than that. In particular, it needs to get started quickly in the late 1800s, and perturbations of entry costs struggle with this feature.

6 Trade: Firm Dynamics without Special Generation

In this section, we return to the stationary age distribution observed among top retailers/wholesalers, which resembles the features of the baseline model in Section 4. One natural question is why, unlike industrials, a special generation did not emerge among top retailers and wholesalers. In Section 5.1 we find that both rapid technological changes and the presence of learning are necessary for the emergence of the special generation among industrials. The absence of shocks comparable to the Second Industrial Revolution could contribute to the lack of a special generation. But even with such shocks, the absence of learning would make a special generation more difficult to sustain.

With the help of the model, we find that learning seems weak in retail/wholesale. In particular, positive learning generates too many old firms among the top ones, relative to what we observe in the data. To reach this conclusion, we examine two calibrations. The first imposes no learning, $\zeta = 0$, and we examine comparative statics for the learning rate. The second introduces the learning rate as an additional parameter with $\zeta \geq 0$, disciplined by median age of top firms.

Calibration We calibrate the baseline model in Section 4 to data moments in retail/wholesale. This model is a special case of those we calibrated in Section 5, with no modern technology, and no time-varying market size or entry costs. The 2018 *Fortune* 1,000 list includes 159 retailers and wholesalers, which represent around 0.019% of firms in this sector in the Census BDS data. Accordingly, we use the top 0.019% to represent these top firms in the model.

In the first calibration, we consider the case with no learning (i.e., ζ= 0). Here we only need the first nine parameters in column (1) of Table 2, Panel A, which govern the stationary equilibrium of the baseline model. Correspondingly, we use the first nine data moments in column (1) of Table 2, Panel B, which primarily inform these parameters. We measure the data moments using the Census Bureau's Revenue LBD like in Section 5, but now for the

Table 3: Calibration for Retail/Wholesale Trade

Panel A shows the calibrated parameters for the retail/wholesale trade sector in the baseline model without and with learning. The initial stochastic component of productivity, s_{i0}, is drawn from a distribution F, where $F \sim \text{LogNormal}(\mu_F, \sigma_F^2)$. The permanent component of productivity, e_i, is drawn from an adjusted lognormal distribution $G(e)$ that includes a superstar state to increase fat-tailedness (Internet Appendix IA2.1). Externally calibrated parameters are the same as Table 1. Panel B shows a comparison of data moments and model moments. Growth rates, exit rates, and top firm revenue shares are averaged over the 2014 to 2018 period using Census Revenue LBD data. Top firms are the top 0.019% by revenue. Revenue LBD data moments calculated on FSRDC Project Number 1731. (CBDRB-FY25-P1731-R12291)

Panel A. Internally Calibrated Parameters

Definition **(1)**	**Parameter** **(2)**	**Without Learning** **(3)**	**With Learning** **(4)**
Entry cost	c_e	33.957	33.457
Overhead cost	c_f	15.870	16.044
Persistence of AR(1)	ρ	0.918	0.946
Volatility of AR(1)	σ_ε	0.236	0.201
Initial condition mean	μ_F	-0.464	-0.195
Initial condition dispersion	σ_F	0.621	0.527
Permanent type dispersion	σ_G	0.142	0.140
Superstar probability	p_G^*	4e-5	4e-5
Superstar factor	λ_G^*	8.264	8.532
Learning rate	ζ	—	2e-6

Panel B. Calibration Moments

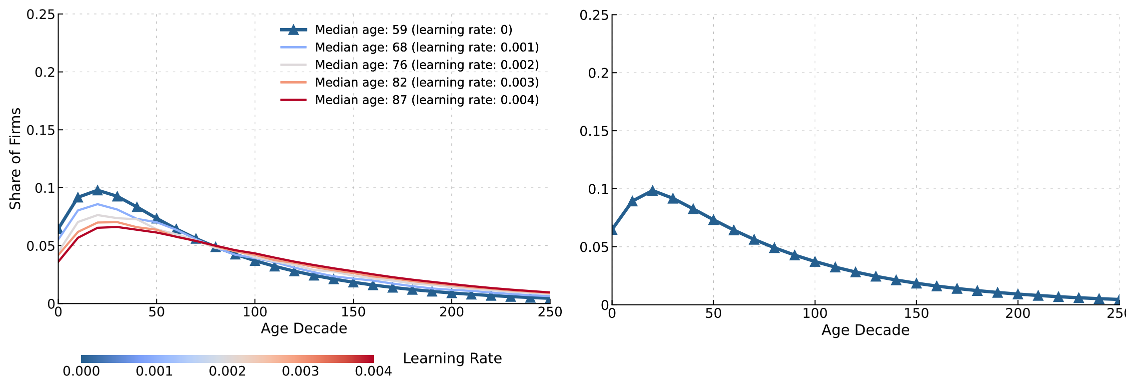
Moment **(1)**	**Data** **(2)**	**Without Learning** **(3)**	**With Learning** **(4)**
Age 1–5 revenue growth rate	0.001	0.001	0.000
Age 6–10 revenue growth rate	-0.018	-0.016	-0.014
Age 11–15 revenue growth rate	-0.020	-0.021	-0.023
Age 1–2 exit rate	0.152	0.152	0.163
Age 3–5 exit rate	0.110	0.096	0.104
Age 6–10 exit rate	0.084	0.072	0.078
Age 11–15 exit rate	0.069	0.058	0.065
Top 1% entrants revenue share	0.253	0.256	0.257
Top firms revenue share	0.390	0.391	0.387
Top firms median age	59	—	59

retail/wholesale sector (again averaging over 2014 to 2018). We continue to use the same externally calibrated parameters.[23] In the second calibration, we allow for learning. In this case, the learning rate ζ is the tenth parameter. We discipline the learning rate by introducing the median age of top retailers and wholesalers in 2018 as an additional moment. As we discuss

[23]The exit rate of large firms, which we used to discipline the exogenous exit rate, is similar across the retail/wholesale and industrial sectors according to the Census Business Dynamic Statistics dataset.

Figure 12: Stationary Age Distribution for Top Retailers/Wholesalers

Panel A shows the model-implied age decade distribution for the top 0.019% retailers and wholesalers in 2018. The solid blue line with triangles represents the benchmark case without learning, using the parameters in column (3) of Table 3, Panel A. The other lines illustrate how the distribution changes if the learning rate increases up to 0.004, with colors transitioning from cool to warm. Panel B is the model-implied age decade distribution of the top 0.019% retailers and wholesalers with calibrated learning rate. The model uses the parameters in column (4) of Table 3, Panel A.



Panel A. Baseline Calibration

Panel B. Calibration with Learning

below, the calibrated learning rate is essentially zero.

In both cases, we calibrate the parameters to match moments of the stationary equilibrium to their data counterparts. Columns (3) and (4) of Table 3, Panel A, show the calibrated parameters. The calibrated parameters in these two cases are extremely similar, and the calibrated learning rate in the second case is basically zero. In other words, the model favors no learning for the retail/wholesale sector. As before, the calibration generally requires high dispersion of the permanent type and the initial condition to match the top entrant revenue share.

Results Figure 12 plots the age distribution among top firms, which is a key pattern in the data (Figure 6) that we do not explicitly target in the calibration. In Panel A, the solid blue line with triangles shows top firms' age distribution in the first calibration with no learning. Although this calibration does not target any moment related to firm age, the median age is 59 just like in the data. The additional lines in Panel A shows the age distribution of top firms if we increase the learning rate from $\zeta = 0$ to 0.004, holding all other parameters constant. We see that, with higher learning rates, older firms accumulate more experience and are less likely to exit, and the median age becomes too high. In Panel B, the solid blue line with triangles shows top firms' age distribution in the second calibration. The median age in the model is again 59. The learning rate we obtain is barely above zero. As mentioned above, the dispersion of the permanent component of productivity is high to match the top entrant revenue share. With a high permanent type, accumulating a sufficiently negative stochastic component to exit is less likely. Increasing the learning rate only further reduces the probability of endogenous exit, which would make top firms too old.

Overall, these analyses suggest that learning appears weak in the retail/wholesale sector. As a result, even if firms in a given cohort obtain technological or other advantages, they are more easily taken over by subsequent entrants, and a special generation would be difficult to sustain.

7 Conclusion

We study the largest American companies at different points in time and document several patterns. In industrials, we observe the persistent prominence of firms born around 1900, even though the particular companies have changed substantially over time. As a result, the average age of the largest industrials keeps increasing. In retail and wholesale, we observe the age distribution of the largest firms to be stationary. In services, very large firms became more common after the 1970s and the 1980s. These decades could have produced a special generation of entrants similar to the experience of industrials in the early 20th century, but this possibility will take more decades to verify.

The evidence suggests that certain settings produce special generations of entrants that give rise to superstar firms for decades to come, but these settings occur occasionally. Through the lenses of our model analyses, these settings occur if new technologies emerge that exhibit economies of scale, confer low adoption costs for new entrants, and require organizational learning. The combination of these forces leads to "hysterisis," and produces special cohorts that have a strong edge relative to both firms that came before and potential entrants thereafter. At the same time, the individual firms at the top can keep churning due to idiosyncratic shocks, so top firms are old yet the landscape is not stale. Taken together, our facts and analyses provide new perspectives on the dynamics of top firms.

References

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. Innovation, reallocation, and growth. *American Economic Review*, 2018. 108(11):3450–91.

Aghion, Philippe and Peter Howitt. A model of growth through creative destruction. *Econometrica*, 1992. 60(2):323–351.

Aghion, Philippe and Jean Tirole. The management of innovation. *Quarterly Journal of Economics*, 1994. 109(4):1185–1209.

Arrow, Kenneth J. *The Limits of Organization*. W. W. Norton & Company, 1974.

Atkeson, Andrew and Patrick J Kehoe. Modeling the transition to a new economy: Lessons from two technological revolutions. *American Economic Review*, 2007. 97(1):64–88.

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.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo, and Jonathan Timmis. Coverage and representativeness of Orbis data. 2020.

Beraja, Martin and Francisco Buera. The life-cycle of concentrated industries. Working paper, 2024.

Berle, Adolf and Gardiner Means. *The Modern Corporation and Private Property*. Harcourt, Brace & Company, 1932.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. Belief overreaction and stock market puzzles. *Journal of Political Economy*, 2024. 132(5):1450–1484.

Braguinsky, Serguey, Joonkyu Choi, Yuheng Ding, Karam Jo, and Seula Kim. Megafirms and recent trends in the US innovation: Empirical evidence from the US patent data. Working paper, 2024.

Buera, Francisco J. and Yongseok Shin. Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy*, 2013. 121(2):221–272. URL http://www.jstor.org/stable/10.1086/670271.

Carroll, Glenn R and Michael T Hannan. *The Demography of Corporations and Industries*. Princeton University Press, 2000.

Carter, Susan B, Scott Sigmund Gartner, Michael R Haines, Alan L Olmstead, Richard Sutch, and Gavin Wright. *Historical Statistics of the United States*, volume 3. Cambridge University Press New York, 2006.

Chandler, Alfred D. Organizational capabilities and the economic history of the industrial enterprise. *Journal of Economic Perspectives*, 1992. 6(3):79–100.

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

Christensen, Clayton M. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business School Press, 1997.

Clementi, Gian Luca and Berardino Palazzo. Entry, exit, firm dynamics, and aggregate fluctuations. *American Economic Journal: Macroeconomics*, 2016. 8(3):1–41.

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.

Crouzet, Nicolas and Neil R. Mehrotra. Small and large firms over the business cycle. *American Economic Review*, 2020. 110(11):3549–3601.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda. Where has all the skewness gone? The decline in high-growth (young) firms in the US. *European Economic Review*, 2016. 86:4–23.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda. Changing business dynamism and productivity: Shocks versus responsiveness. *American Economic Review*, 2020. 110(12):3952–3990.

Doidge, Craig, G Andrew Karolyi, and René M Stulz. The US listing gap. *Journal of Financial Economics*, 2017. 123(3):464–487.

Draghi, Mario. The future of European competitiveness - a competitiveness strategy for Europe. 2024. URL https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en.

Fama, Eugene F and Kenneth R French. The cross-section of expected stock returns. *Journal of Finance*, 1992. 47(2):427–465.

Gabaix, Xavier. The granular origins of aggregate fluctuations. *Econometrica*, 2011. 79(3):733–772.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow. How destructive is innovation? *Econometrica*, 2019. 87(5):1507–1541.

Gordon, Robert J. Does the "new economy" measure up to the great inventions of the past? *Journal of Economic Perspectives*, 2000. 14(4):49–74.

Gordon, Robert J. *The Rise and Fall of American Growth*. Princeton University Press, 2016.

Guntin, Rafael and Federico Kochen. The origins of top firms. Working paper, 2025.

Gutiérrez, Germán, Callum Jones, and Thomas Philippon. Entry costs and aggregate dynamics. *Journal of Monetary Economics*, 2021. 124:S77–S91.

Hannah, Leslie. *The Rise of the Corporate Economy*. Routledge, 1976.

Hannah, Leslie. Survival and size mobility among the world's largest 100 industrial corporations, 1912-1995. *American Economic Review*, 1998. 88(2):62–65.

Harris. Founding dates of the 1994 Fortune 500 US companies. *Business History Review*, 1996. 70(1):69–90.

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.

Hopenhayn, Hugo A. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 1992. 60(5):1127–1150.

Hornbeck, Richard, Shanon Hsuan-Ming Hsu, Anders Humlum, and Martin Rotemberg. Gaining steam: Incumbent lock-in and entrant leapfrogging. Working paper, 2024.

Jovanovic, Boyan. Fitness and age: Review of Carroll and Hannan's "Demography of Corporations and Industries". *Journal of Economic Literature*, 2001. 39(1):105–119.

Jovanovic, Boyan and Glenn M MacDonald. The life cycle of a competitive industry. *Journal of Political Economy*, 1994. 102(2):322–347.

Jovanovic, Boyan and Peter L Rousseau. General purpose technologies. In *Handbook of Economic Growth*, volume 1, pages 1181–1224. Elsevier, 2005.

Kaplan, Steven N, Berk A Sensoy, and Per Strömberg. Should investors bet on the jockey or the horse? Evidence from the evolution of firms from early business plans to public companies. *Journal of Finance*, 2009. 64(1):75–115.

Karahan, Fatih, Benjamin Pugsley, and Ayşegül Şahin. Demographic origins of the start-up deficit. *American Economic Review*, 2024. 114(7):1986–2023.

Klepper, Steven. Entry, exit, growth, and innovation over the product life cycle. *American Economic Review*, 1996. 86(3):562–583.

Klette, Tor Jakob and Samuel Kortum. Innovating firms and aggregate innovation. *Journal of Political Economy*, 2004. 112(5):986–1018.

Kuvshinov, Dmitry and Kaspar Zimmermann. The big bang: Stock market capitalization in the long run. *Journal of Financial Economics*, 2022. 145(2):527–552.

Kwon, Spencer Y, Yueran Ma, and Kaspar Zimmermann. 100 years of rising corporate concentration. *American Economic Review*, 2024. 114(7):2111–2140.

Lamoreaux, Naomi R. *The Great Merger Movement in American Business, 1895-1904*. Cambridge University Press, 1988.

Lee, Yoonsoo and Toshihiko Mukoyama. A model of entry, exit, and plant-level dynamics over the business cycle. *Journal of Economic Dynamics and Control*, 2018. 96:1–25.

Levitt, Steven D, John A List, and Chad Syverson. Toward an understanding of learning by doing: Evidence from an automobile assembly plant. *Journal of Political Economy*, 2013. 121(4):643–681.

Louçã, Francisco and Sandro Mendonça. Steady change: The 200 largest US manufacturing firms throughout the 20th century. *Industrial and Corporate Change*, 2002. 11(4):817–845.

Luttmer, Erzo G. J. On the mechanics of firm growth. *Review of Economic Studies*, 2011. 78(3):1042–1068.

Marshall, Alfred. *Principles of Economics*. Macmillan, London, 8 edition, 1920. 8th edition.

Meeks, Geoffrey and Geoffrey Whittington. Death on the stock exchange: The fate of the 1948 population of large UK quoted companies, 1948–2018. *Business History*, 2023. 65(4):679–698.

Miranda, Mario J. and Paul L. Fackler. *Applied Computational Economics and Finance*. MIT Press, Cambridge, MA, 2002.

Mokyr, Joel. The second industrial revolution, 1870-1914. In *Storia dell'economia Mondiale*, pages 219–245. Laterza Rome, 2000.

Moreira, Sara. Firm dynamics, persistent effects of entry conditions, and business cycles. Working paper, 2017.

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

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter. Plants in space. *Journal of Political Economy*, 2024. 132(3):867–909.

Ohyama, Atsushi, Serguey Braguinsky, and Kevin Murphy. Entrepreneurial ability and market selection in an infant industry: Evidence from the Japanese cotton spinning industry. *Review of Economic Dynamics*, 2004. 7:354–381.

O'Reilly, Charles A and Michael L Tushman. *Lead and Disrupt: How to Solve the Innovator's Dilemma*. Stanford University Press, 2021.

Prais, Sigbert J. *The Evolution of Giant Firms in Britain*. Cambridge University Press Cambridge, 1976.

Prais, Sigbert J and Caroline Reid. Large and small manufacturing enterprises in Europe and America. In *Markets, Corporate Behaviour and the State: International Aspects of Industrial Organization*, pages 78–94. Springer, 1976.

Schlingemann, Frederik P. and René M. Stulz. Have exchange-listed firms become less important for the economy? *Journal of Financial Economics*, 2022. 143(2):927–958.

Schumpeter, Joseph Alois. *Capitalism, Socialism and Democracy*. Allen & Unwin, London, 1942.

Sedláček, Petr and Vincent Sterk. The growth potential of startups over the business cycle. *American Economic Review*, 2017. 107(10):3182–3210.

Singla, Shikhar. Regulatory costs and market power. Working paper, 2023.

Stangler, Dane. The economic future just happened. Working paper, 2009.

Stangler, Dane and Samuel Arbesman. What does Fortune 500 turnover mean? *Available at SSRN 2078162*, 2012.

Sterk, Vincent, Petr Sedláček, and Benjamin Pugsley. The nature of firm growth. *American Economic Review*, 2021. 111(2):547–79.

Tauchen, George. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 1986. 20(2):177–181.

The Economist. America's corporate giants are getting harder to topple. *The Economist*, 2023. URL https://www.economist.com/business/2023/08/21/americas-corporate-giants-are-getting-harder-to-topple.

Internet Appendix: For Online Publication

IA1 Additional Figures and Tables

Figure IA1: Sales by Birth Decade for 2018 *Fortune* 1,000 Companies

This figure shows the total sales of 2018 *Fortune* 1,000 companies by birth decade. Companies are assigned to main sectors based on their industries in 2018. The main sectors correspond to SIC codes 15-17 (construction), 10-14 and 20-39 (industrials), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

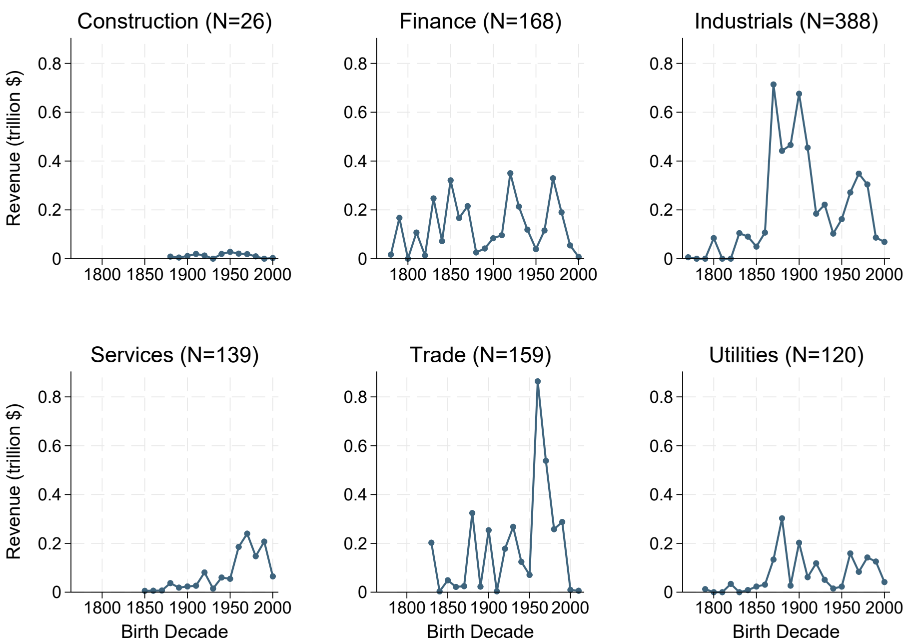


Figure IA2: Employment by Birth Decade for 2018 *Fortune* 1,000 Companies

This figure shows the total employment of 2018 *Fortune* 1,000 companies by birth decade. Companies are assigned to main sectors based on their industries in 2018. The main sectors correspond to SIC codes 15-17 (construction), 10-14 and 20-39 (industrials), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

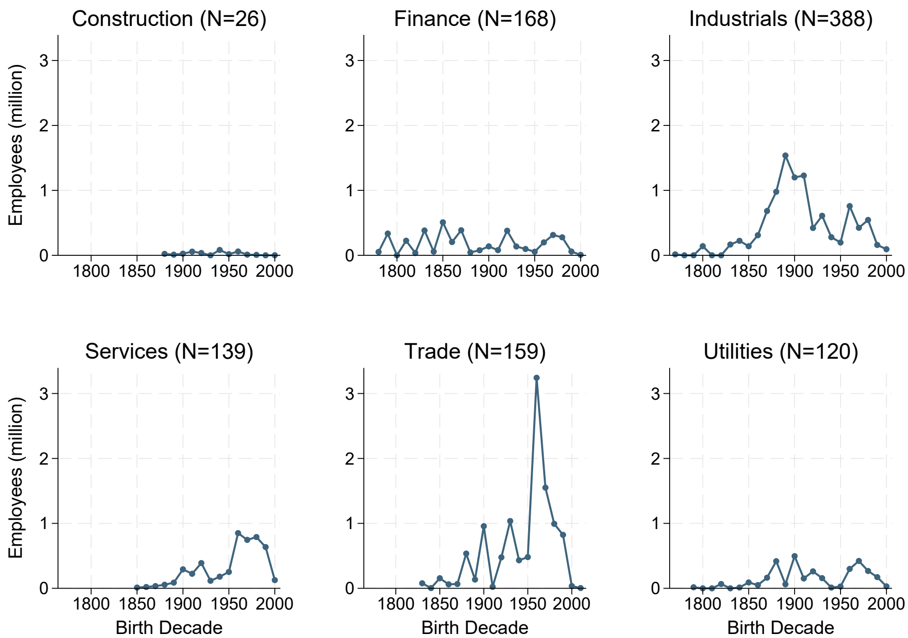


Figure IA3: Top 1,000 in 2018 with Private Firms from *Forbes* List

This figure shows the distribution of birth decade among the largest 1,000 firms by sales in the 2018 *Fortune* 1,000 list (solid line with triangles), and among the combined list with the largest 1,000 firms by sales including additional large private firms according to the 2018 *Forbes: America's Largest Private Companies* list (solid line with circles). The main sectors correspond to SIC codes 15-17 (construction), 10-14 and 20-39 (industrials), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

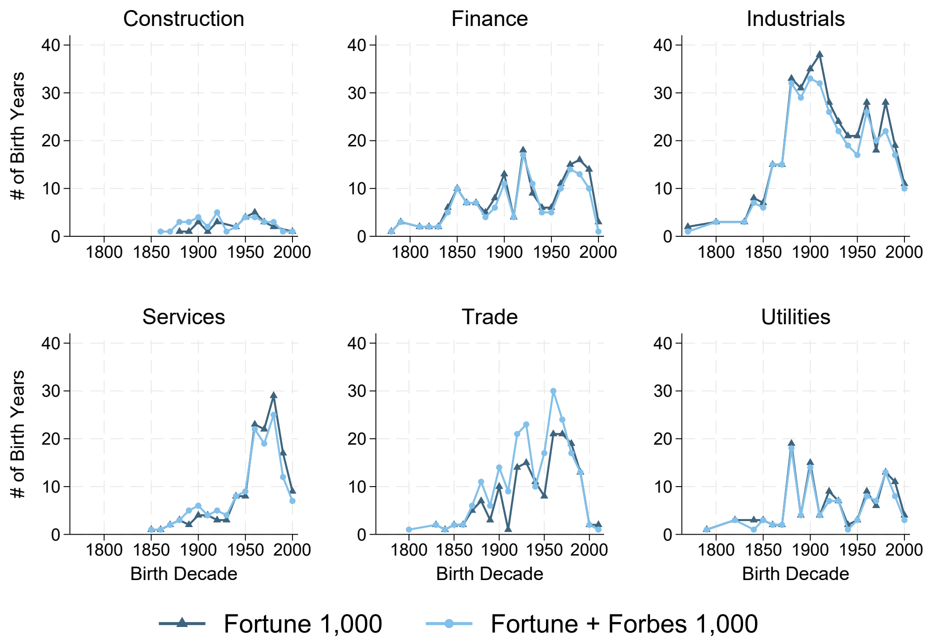


Figure IA4: Birth Year Distribution Excluding Firms with Complex Histories

This figure shows the distribution of birth years per decade for the 2018 *Fortune* 1,000 companies, excluding firms that involve mergers of equals and firms that are spin-offs where we do not know the origin of the spun-off entity. The main sectors correspond to SIC codes 15-17 (construction), 10-14 and 20-39 (industrials), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

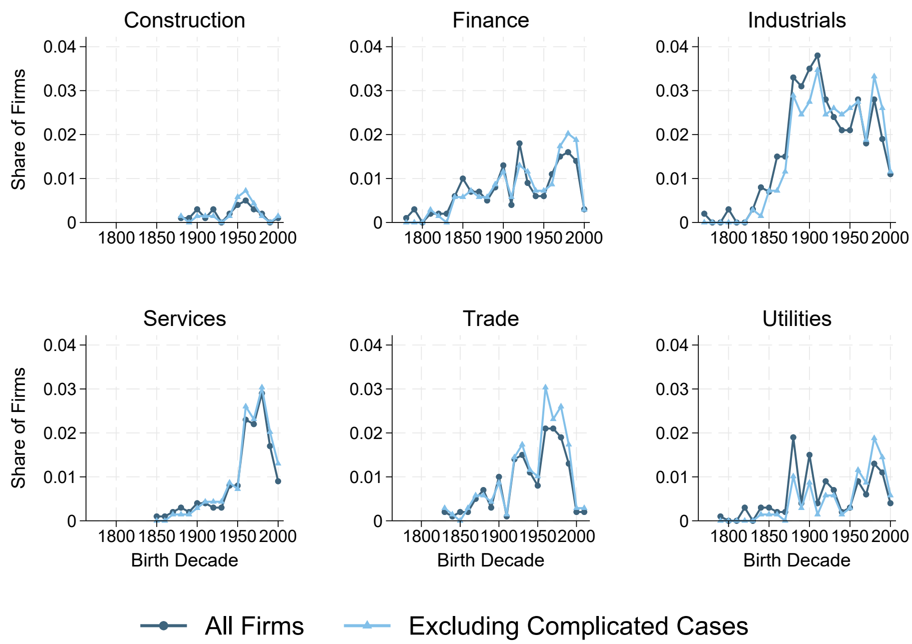


Figure IA5: Firm Age in *Fortune* 1,000 and BDS Firms with More than 10,000 Employees

This figure compares the age distribution among the 2018 *Fortune* 1,000 list and among firms in the BDS with more than 10,000 employees. BDS suppresses the number of firms by age category in some sectors for the very largest firms. We impute values whenever possible and otherwise set the suppressed cells equal to zero. The main sectors correspond to SIC codes 15-17 (construction), 10-14 and 20-39 (industrials), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

Figure IA6: Largest Industrials in 1955 and 2018

The solid line with triangles shows the number of birth years per decade for the largest 388 industrial companies in the 2018 *Fortune* list (which contains 388 industrial firms). The solid line with circles shows the number of birth years per decade for the largest 388 industrial companies in the 1955 *Fortune* list. The dashed line with circles is the solid line shifted to the right for six decades (i.e., what the 2018 distribution would be like if the age distribution of top firms were stationary).

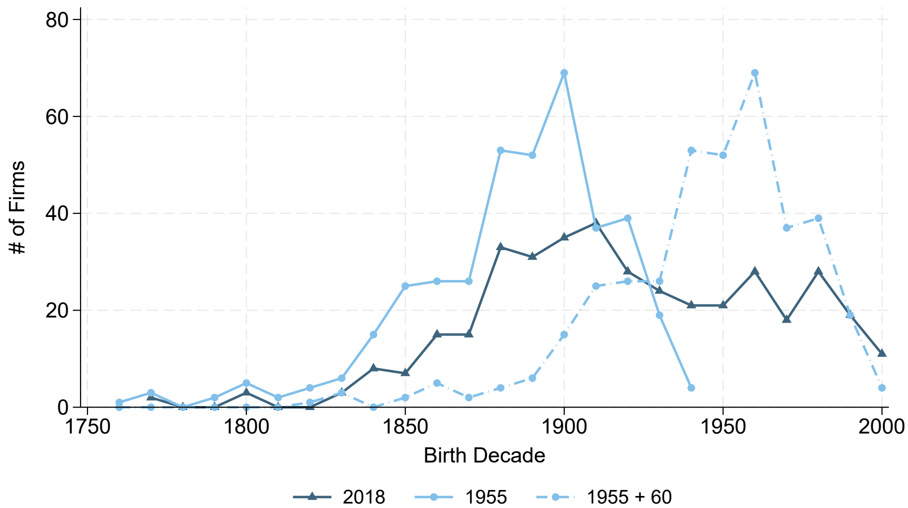


Figure IA7: Birth Years of 1956 *Fortune* Largest 50 Retailers and Wholesalers

This figure shows the number of birth years per decade for the largest 50 merchandising firms (i.e., retailers and wholesalers) in 1956 by sales.

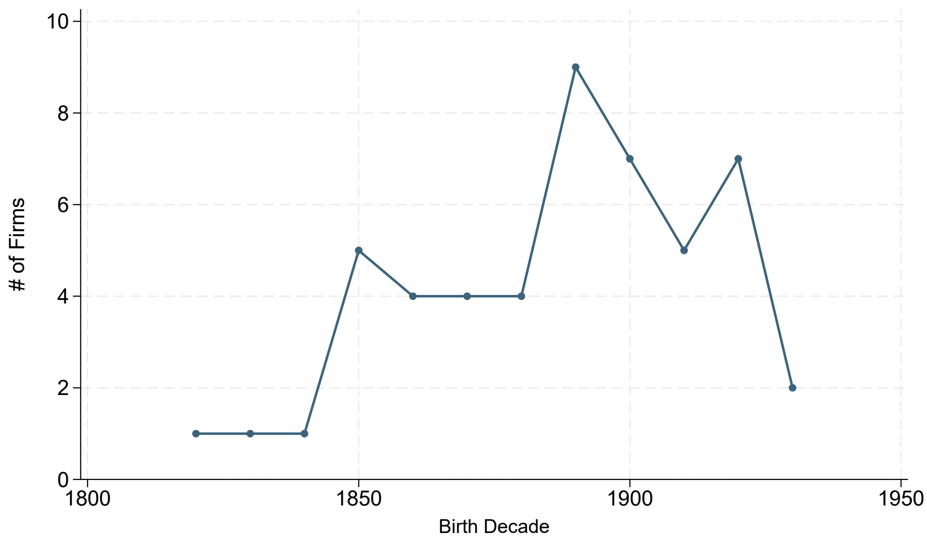


Figure IA8: Sector Composition of the Largest 1,000 Firms

This figure plots the sector composition of the largest 1,000 firms in the U.S. by sales in 1959 and 2018. The 1959 composition is estimated using *Statistics of Income* tabulations of corporations by sales (Kwon, Ma, and Zimmermann, 2024). First, we estimate a threshold for the largest 1,000 firms by sales. We then use sector-level tabulations to interpolate how many firms in each sector are above the threshold. The 2018 composition directly uses the *Fortune* list. The main sectors correspond to SIC codes 15-17 (construction), 10-14 and 20-39 (industrials), 40-49 (transportation, communications, and utilities), 50-59 (wholesale and retail trade), 60-67 (finance, insurance, and real estate), and 70-89 (services).

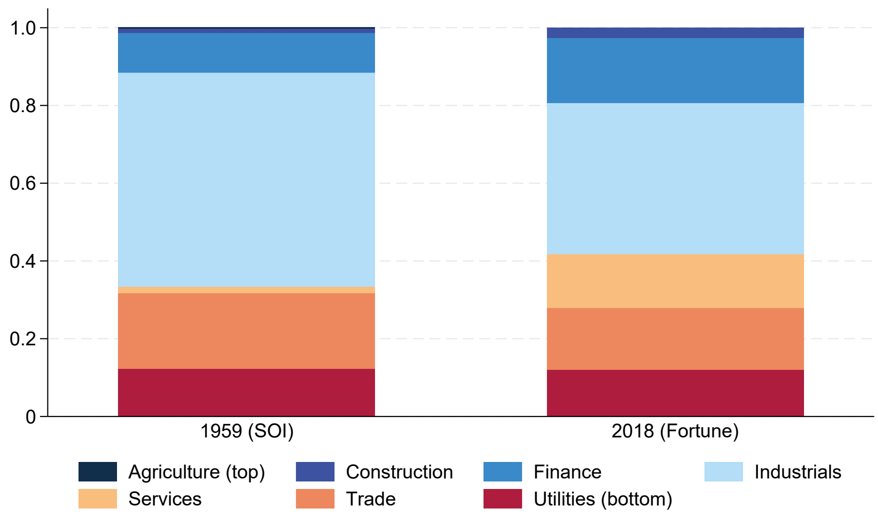


Figure IA9: Modern Technology Shock

This figure plots the perfect foresight path of the modern technology adoption cost realized in 1860. Before 1860, the adoption cost is sufficiently high and no firms adopt the modern technology. The adoption cost reaches its minimum in 1899 and then remains constant. The perfect foresight path follows the parameters in column (3) of Table 2, Panel A.

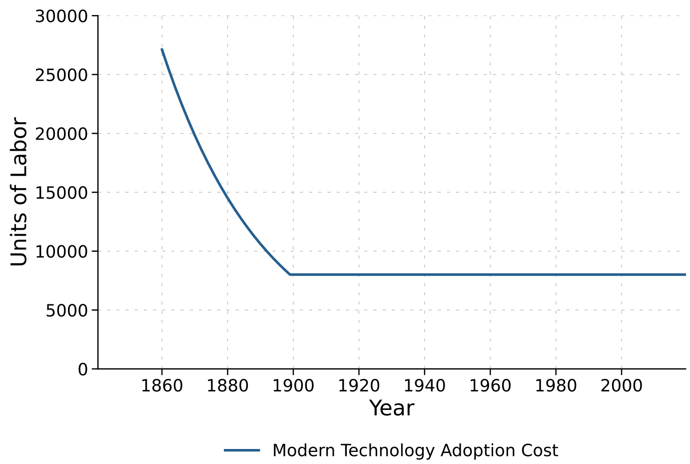
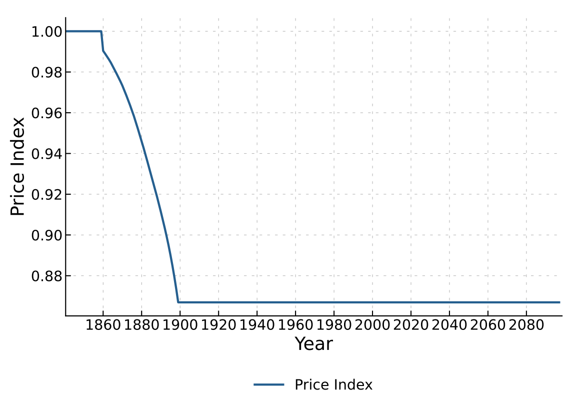
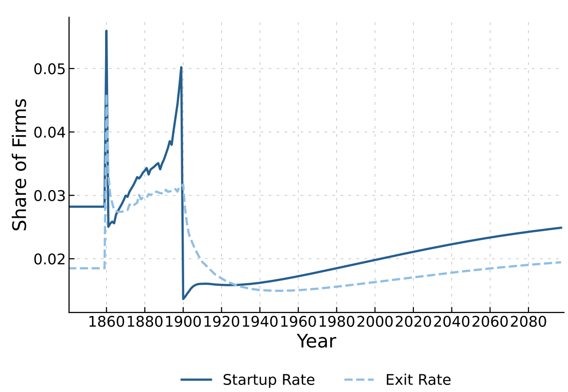


Figure IA10: Price of Output, Startup Rate, and Exit Rate with Modern Technology Shock

Panel A plots the price of output, p_t, normalized to an initial value of one, following the modern technology shock. The adoption cost of the modern technology declines starting 1860, reaches its minimum in 1899, and then remains constant. Panel B plots the startup rate and exit rate over the same period. The model uses parameters in column (3) of Table 2, Panel A.



Panel A. Price of Output



Panel B. Startup Rate and Exit Rate

Figure IA11: Share of Firms using Modern Technology

This figure plots the share of entrants (left y-axis) and all firms (right y-axis) using the modern technology. The adoption cost of the modern technology declines starting 1860, reaches its minimum in 1899, and then remains constant. The model uses parameters in column (3) of Table 2, Panel A.

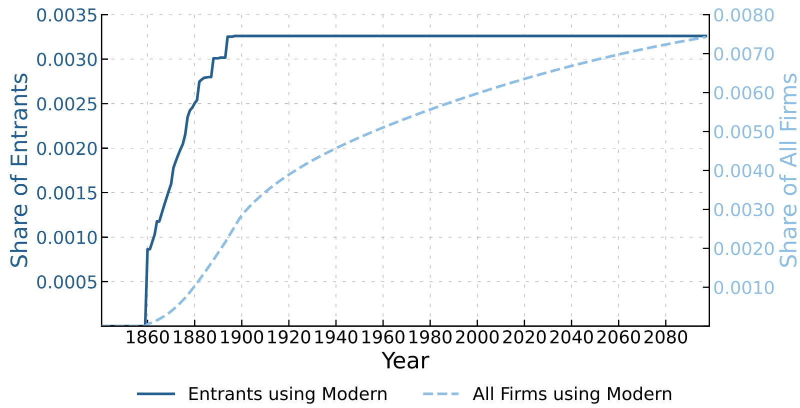


Figure IA12: Convergence to Equilibrium following Modern Technology Shock

This figure plots the age decade distribution of the top 0.15% industrial firms in 2010, 2110, 2360, and 2860, corresponding to 150, 250, 500, and 1000 years after the beginning of the modern technology shock (i.e., long-horizon model solution counterpart to Figure 3). The adoption cost of the modern technology declines starting 1860, reaches its minimum in 1899, and then remains constant. The model uses parameters in column (3) of Table 2, Panel A.

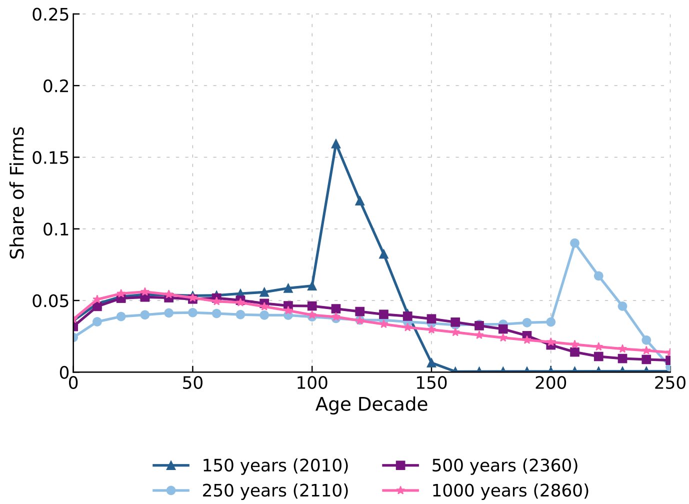
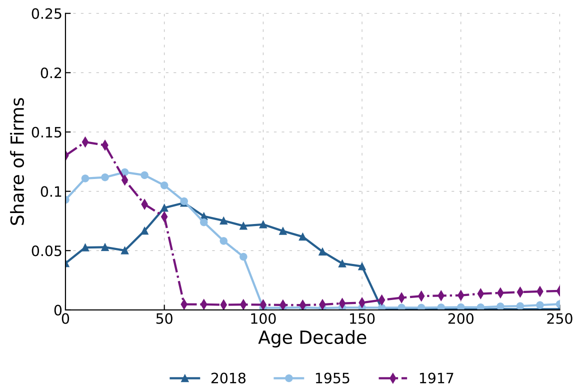
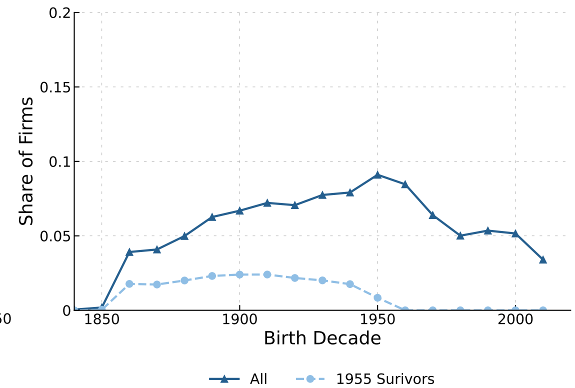


Figure IA13: Top Firms in Model with Slow Decline of Modern Technology Adoption Cost

Panel A is the model-implied age decade distribution of the top 0.15% industrial firms in 2018, 1955, and 1917. Panel B is the model-implied birth decade distribution of the top 0.15% industrial firms in 2018 and the subset that were also among the top 0.15% industrial firms in 1955. The model uses parameters in column (3) of Table 2, Panel A, except that the adoption cost decreases exponentially for 100 periods starting 1860. The terminal adoption cost, c_{M1959}, is set to the same value as c_{M1899} in column (3) of Table 2, Panel A. We simulate the model 500 times, each with an initial 100K firms drawn from the ergodic distribution, a 500-period burn-in, and then for 160 periods following the shock. Results are averaged over 500 simulations.



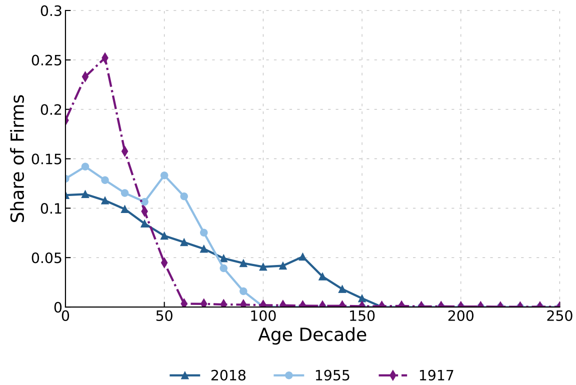
Panel A. Age Distribution of Top Firms



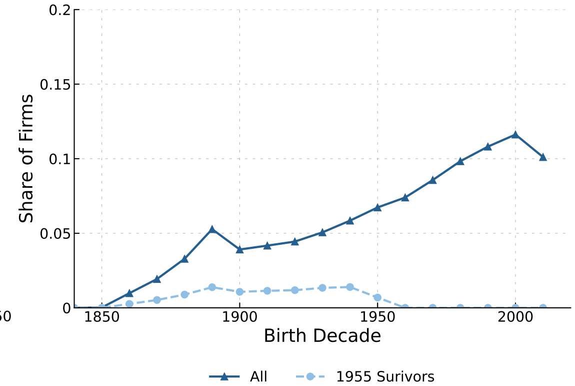
Panel B. Top Firms in 2018 and 1955

Figure IA14: Top Firms in Model with Modern Technology Shock and No Learning

Panel A is the model-implied age decade distribution of the top 0.15% industrial firms in 2018, 1955, and 1917. Panel B is the model-implied birth decade distribution of the top 0.15% industrial firms in 2018 and the subset that were also among the top 0.15% industrial firms in 1955. The model uses parameters in column (3) of Table 2, Panel A, except that the learning rate ζ is 0. We simulate the model 500 times, each with an initial 100K firms drawn from the ergodic distribution, a 500-period burn-in, and then for 160 periods following the shock. Results are averaged over 500 simulations.



Panel A. Age Distribution of Top Firms



Panel B. Top Firms in 2018 and 1955

Figure IA15: Market Size Growth Shock

This figure plots the perfect foresight path of the market size growth rate realized in 1860. The market size growth rate is high between 1860 and 1969, falls in 1970, and then stays constant. The perfect foresight shock follows parameters in column (4) of Table 2, Panel A.

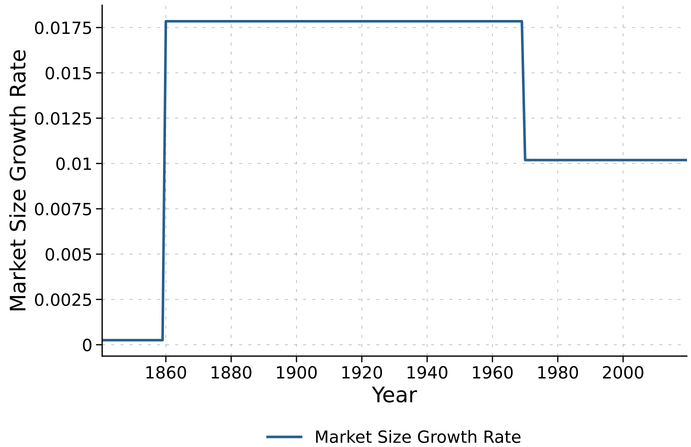
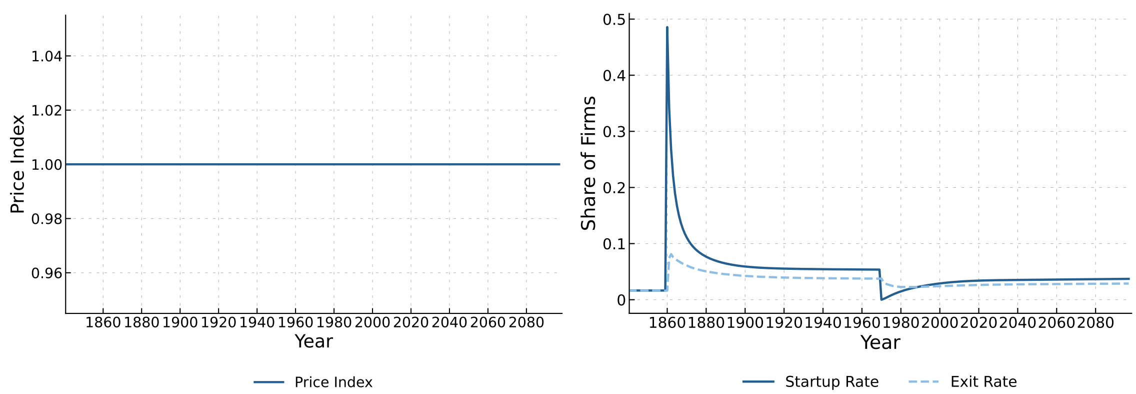


Figure IA16: Price of Output, Startup Rate, and Exit Rate with Market Size Growth Shock

Panel A plots the price index for the price of output, p_t, normalized to an initial value of one, following the market size growth shock. The market size growth rate is high between 1860 and 1969, falls in 1970, and then stays constant. Panel B plots the startup rate and exit rate over the same period. The model uses parameters in column (4) of Table 2, Panel A.



Panel A. Price of Output

Panel B. Startup Rate and Exit Rate

Figure IA17: Entry Cost Shock

This figure plots the perfect foresight path of the entry cost realized in 1860. The entry cost starts rising in 1860, reaches its maximum in 2061, and then remains constant. The perfect foresight shock follows parameters in column (5) of Table 2, Panel A.

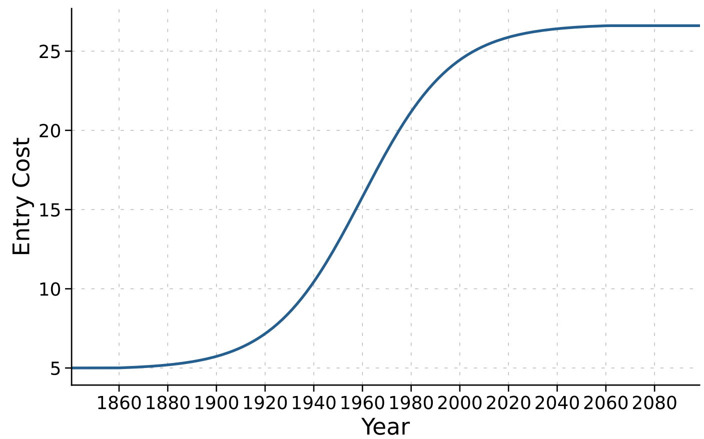
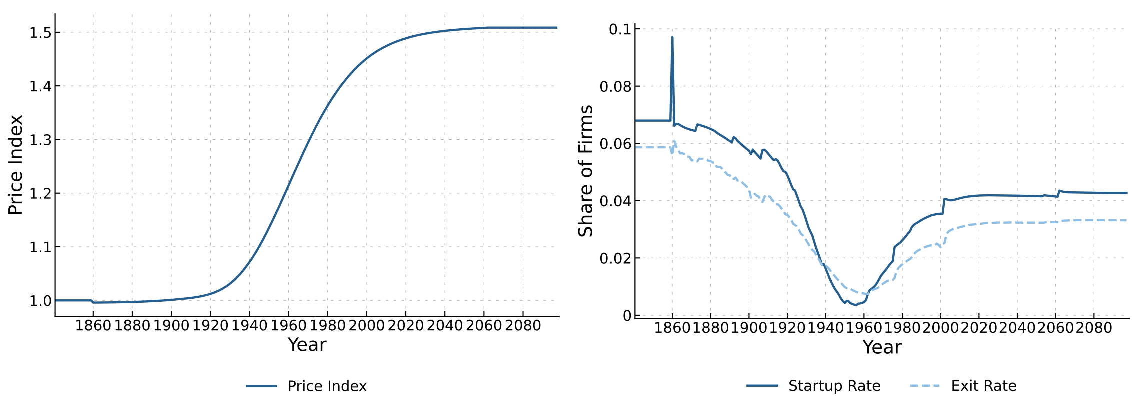


Figure IA18: Price of Output, Startup Rate, and Exit Rate with Entry Cost Shock

Panel A plots the price index for the price of output, p_t, normalized to an initial value of one, following the time-varying entry cost shock. The entry cost starts rising in 1860, reaches its maximum in 2061, and then remains constant. Panel B plots the startup rate and exit rate over the same period. The model uses parameters in column (5) of Table 2, Panel A.



Panel A. Price of Output

Panel B. Startup Rate and Exit Rate

Table IA1: Persistence among the Largest Industrials

This table shows the number of firms by birth decade for the top 388 industrials in 1995 and 2018. "Direct survivors" are top firms in 1955 that are still among the top firms in 2018. "Indirect survivors" are top firms in 1955 that have been acquired by a top firm in 2018.

2018 List		1955 List		Direct Survivor		Direct & Indirect Survivor	
Birth Decade (1)	Number (2)	Birth Decade (3)	Number (4)	Birth Decade (5)	Number (6)	Birth Decade (7)	Number (8)
1760	0	1760	1	1760	0	1760	0
1770	2	1770	3	1770	1	1770	2
1780	0	1780	0	1780	0	1780	0
1790	0	1790	2	1790	0	1790	1
1800	3	1800	5	1800	2	1800	2
1810	0	1810	2	1810	0	1810	0
1820	0	1820	4	1820	0	1820	1
1830	3	1830	5	1830	3	1830	4
1840	8	1840	15	1840	3	1840	6
1850	7	1850	24	1850	3	1850	8
1860	15	1860	26	1860	4	1860	9
1870	15	1870	26	1870	9	1870	13
1880	33	1880	53	1880	12	1880	23
1890	31	1890	51	1890	16	1890	27
1900	35	1900	68	1900	13	1900	25
1910	38	1910	37	1910	9	1910	17
1920	28	1920	38	1920	3	1920	15
1930	24	1930	19	1930	3	1930	9
1940	21	1940	4	1940	0	1940	1
1950	21	1950		1950		1950	
1960	28	1960		1960		1960	
1970	18	1970		1970		1970	
1980	28	1980		1980		1980	
1990	19	1990		1990		1990	
2000	11	2000		2000		2000	

Table IA2: Examples of Turnover among Top Firms

This table lists top ten firms in food (Panel A) and metals (Panel B) manufacturing in 1955 and 2018 and their birth years. It also includes information about the outcomes of the top firms in 1955 by 2018. Outcomes with direct connections to 2018 *Fortune* 1,000 firms are displayed in bold, and outcomes with indirect connections to 2018 *Fortune* 1,000 firms are displayed in italics.

Panel A. Top Ten in Food: 1955 vs 2018

Top in 1955	**Birth Year**	**Outcome in 2018**	**Top in 2018**	**Birth Year**
Swift	**1855**	acquired, owned by JBS	PepsiCo	**1898**
Armour	**1867**	dissolved	Archer Daniels Midland	**1902**
National Dairy Products	**1898**	**Mondelez**	Tyson Foods	**1935**
General Foods	**1895**	*acquired, owned by Mondelez*	Coca-Cola	**1886**
Borden	**1857**	dissolved	Kraft Heinz	**1869**
Wilson & Co.	**1855**	dissolved	Mondelez	**1898**
General Mills	**1866**	**General Mills**	General Mills	**1866**
Cudahy Packing	**1887**	acquired, owned by Sigma Alimentos	Land O'Lakes	**1921**
Standard Brands	**1852**	*acquired, owned by Mondelez*	Kellogg	**1906**
Ralston Purina	**1894**	dissolved	Molson Coors Brewing	**1873**

Panel B. Top Ten in Metals: 1955 vs 2018

Top in 1955	**Birth Year**	**Outcome in 2018**	**Top in 2018**	**Birth Year**
U.S. Steel	**1872**	**U.S. Steel**	Nucor	**1905**
Bethlehem Steel	**1857**	dissolved	Arconic	**1888**
Republic Steel	**1886**	acquired, owned by Grupo Simec	Stanley Black & Decker	**1843**
Aluminum Co. of America	**1888**	**Alcoa, Arconic**	U.S. Steel	**1872**
American Can	**1901**	Primerica	Parker-Hannifin	**1917**
Continental Can	**1904**	dissolved	Alcoa	**1888**
Inland Steel	**1893**	acquired, owned by ArcelorMittal	Ball	**1880**
Armco Steel	**1899**	**AK Steel Holdings**	Corning	**1851**
American Metal Products	**1887**	*acquired, owned by Freeport-McMoRan*	Steel Dynamics	**1993**
Jones & Laughlin Steel	**1852**	acquired, owned by Grupo Simec	Crown Holdings	**1892**

IA2 Computational Appendix

IA2.1 State Variables and Approximations

This section describes the method used to construct three state variables in the model: the permanent component of productivity e_i, the stochastic component of productivity s_{iat}, and firm age a_{it}. The same method applies to the models described in Sections 4, 5.1, 5.2, 5.3, and 6; Section 5.1 also incorporates a technology choice that necessitates tracking a traditional versus modern technology type at the firm level.

The permanent component of productivity Permanent component, e_i, is drawn at birth from a modified lognormal distribution, G. We start by approximating a log normal distribution with variance σ_G^2 using three evenly spaced grid points in logs over the range $[-2.5\sigma_G, 2.5\sigma_G]$. The probability of each grid point is initially assigned by evaluating the normal CDF of $N\left(0, \sigma_G^2\right)$ at midpoint intervals. We then introduce an additional fourth point to account for fat-tailedness. This extra point is positioned at $\lambda_G^* \times 2.5\sigma_G$, and is drawn with probability p_G^*. We rescale the probabilities of the original three grid points by a factor of $1 - p_G^*$, i.e., with probability $1 - p_G^*$ the type is drawn from three approximately log normal grid points. The transition matrix for e_i is an identity matrix, because the type remains unchanged over each firm's lifetime.

The stochastic component of productivity Stochastic component, s_{iat}, evolves as an AR(1) process in logs over a firm's lifetime, i.e., $\log s_{iat} = \rho \log s_{ia-1t-1} + \sigma_\varepsilon \varepsilon_{it}$, with initial condition s_{i0} drawn at birth from distribution $F \sim \mathrm{LogNormal}(\mu_F, \sigma_F^2)$. We approximate the AR(1) process using a Tauchen (1986) Markov chain approximation. We use 71 evenly spaced grid points in logs over the interval $\left[-3\sqrt{\frac{\sigma_\varepsilon^2}{1-\rho^2}}, \quad 3\sqrt{\frac{\sigma_\varepsilon^2}{1-\rho^2}}\right]$. The transition matrix for s_{iat} captures the conditional distribution of future stochastic component states given the current state. For the distribution F over the initial condition, we assign the probability over each of the 71 grid points by evaluating the normal CDF of $N\left(\mu_F, \sigma_F^2\right)$ at midpoint intervals.

Firm age Firm age a_{it} begins at zero for new entrants and increases by one per year until a maximum age $\bar{a}$ of 270. Our results are nearly identical for larger maximum ages, e.g. $\bar{a} = 500$. Once the firm age reaches the maximum age of 270, it remains at that age permanently. As a result, the firm age grid is a column vector of consecutive integers ranging from 0 to 270. The corresponding entry probability is one for age 0 and zero for all other ages. The firm age transition matrix is an upper shift identity matrix that increments firm age by one year, with the last row of the transition matrix containing a single one in the final position.

Joint state variable grid We generate a joint state variable grid as the tensor product of each of the three state variables, where each grid point is a unique triple of the three states. The grid has dimensions $76,964 \times 3$, calculated as $4 \times 71 \times 271 = 76,964$. The value function V, policy functions n and x, and firm measure $\bar{\mu}$ at a given time point in the equilibrium solution are represented as column vectors of size $76,964 \times 1$, evaluated at each grid point.

Transition matrix and entry probability The $76,964 \times 76,964$ transition matrix over the triples is the Kronecker product of the transition of each of the three states. Similarly, the $76,964 \times 1$ initial distribution over the triples is the Kronecker of the three initial distributions. Because age and permanent type are deterministic, most of the transition matrix and even initial distribution are empty, and we use sparse matrices for both.

Incorporating the technology choice With the introduction of modern technology in Section 5.1, the joint state variable grid doubles in size to account for firms' two technology choices. Each vector stacks with its counterpart, forming a single expanded column vector: V with V^M, n with n^M, x with x^M, and $\bar{\mu}$ with $\bar{\mu}^M$. As a result, the dimension of each column vector increases to $153,928 \times 1$, where $4 \times 71 \times 2 \times 271 = 153,928$. Accordingly, the transition matrix for the expanded joint grid has dimensions $153,928 \times 153,928$. The entry probability vector also expands to $153,928 \times 1$ to incorporate technology choice for entrants. Grid points corresponding to unselected technologies have an entry probability of zero.

IA2.2 Solving the Stationary Equilibrium

This section describes the method used to solve for the stationary equilibrium. The same method applies to the initial and terminal equilibria described in Sections 5.1, 5.2, and 5.3, as well as the baseline model in Section 4 and 6. The value function, V, policy functions, n, x, and firm measure, $\bar{\mu}$, for traditional technology are computed in all model variants. The counterparts associated with modern technology, V^M, n^M, x^M, $\bar{\mu}^M$, and technology choice, m are solved only in Section 5.1.

In the steady state, we use a recursive method to obtain value functions V and V^M, for all age groups. Given a constant price, we first solve for the firm's value at the maximum age, $V(\bar{a}, e_i, s_{i,\bar{a}})$ and $V^M(\bar{a}, e_i, s_{i,\bar{a}})$, as the fixed point of (2) and (7). We compute the fixed points using value function iteration, with convergence achieved once the maximum norm is below 10^{-7}.

Within the Bellman equations, the flow utility is determined by solving a static optimization problem. Labor demand $n\left(\bar{a}, e_i, s_{i,\bar{a}}\right)$ and $n^M\left(\bar{a}, e_i, s_{i,\bar{a}}\right)$ have closed-form solutions. The expected continuation value is obtained by multiplying the transition matrix with $V\left(\bar{a}, e_i, s_{i,\bar{a}}\right)$ and $V^M\left(\bar{a}, e_i, s_{i,\bar{a}}\right)$, since further aging no longer affects the value function. We incorporate adjustments for the exogenous exit rate and discount factor to compute the discounted expected continuation value. After obtaining these two components, we compute the exit decisions $x\left(\bar{a}, e_i, s_{i,\bar{a}}\right)$ and $x^M\left(\bar{a}, e_i, s_{i,\bar{a}}\right)$.

Once the value function for the oldest firms is determined, the Bellman equations are applied recursively to solve for the value functions, $V\left(a, e_i, s_{i,a}\right)$ and $V^M\left(a, e_i, s_{i,a}\right)$ of younger firms down to the entrants. While solving for $V\left(a, e_i, s_{i,a}\right)$ and $V^M\left(a, e_i, s_{i,a}\right)$, the flow utility remains a static optimization problem, but the firm discounts the value function evaluated at one year older, $V\left(a+1, e_i, s_{i,a+1}\right)$ and $V^M\left(a+1, e_i, s_{i,a+1}\right)$. The technology adoption policy, m, and the expected value of entry are determined after solving for $V\left(0, e_i, s_{i,0}\right)$ and $V^M\left(0, e_i, s_{i,0}\right)$ for entrants. The labor policy functions, n, n^M, and exit policy functions, x, x^M, are determined while computing V and V^M.

The equilibrium price p is determined by minimizing an objective function in which the input is the constant price and the output is the difference between the expected value of entry and the entry cost c_e. This formulation ensures that the free entry condition (3) or (8) holds, so that the expected value of entering the market equals the entry cost. To compute p, we employ the `Broyden` optimization function developed by Miranda and Fackler (2002), with convergence achieved once the maximum norm is below 10^{-12}. Once the equilibrium price is determined, $\bar{\mu}$, $\bar{\mu}^M$, $\bar{M}$ are jointly solved using the product market clearing condition (5) or (11) and the laws of motion (9) and (10).

IA2.3 Solving the Nonstationary Equilibrium

This section describes the method used to solve for the nonstationary equilibrium. The same method applies to the nonstationary equilibria described in Sections 5.1, 5.2, and 5.3. The value function, $\{V_t\}_{t=1}^{T-1}$, policy function, $\{n_t, x_t\}_{t=1}^{T-1}$, firm measure, $\{\bar{\mu}_t\}_{t=1}^{T-1}$, for traditional technology are computed in all model variants. The counterparts associated with modern technology, $\left\{V_t^M, n_t^M, x_t^M, \bar{\mu}_t^M\right\}_{t=1}^{T-1}$, and technology choice, $\{m_t\}_{t=1}^{T-1}$, are solved only in Section 5.1.

To solve the nonstationary equilibrium, we first compute the stationary equilibrium in both the initial and terminal steady states, $\left\{V_0, V_0^M\right\}$, $\left\{n_0, x_0, n_0^M, x_0^M, m_0\right\}$, $\left\{V_T, V_T^M\right\}$, $\left\{n_T, x_T, n_T^M, x_T^M, m_T\right\}$. Starting from the terminal steady state, we solve backward for the value and policy functions $\left\{V_t, V_t^M\right\}_{t=1}^{T-1}$, $\left\{n_t, x_t, n_t^M, x_t^M, m_t\right\}_{t=1}^{T-1}$, using the Bellman equations (2) and (7).

Unlike in the stationary equilibrium, firms discount the value function evaluated both at one year older and one year later, $V_{t+1}\left(a+1, e_i, s_{i,a+1,t+1}\right)$ and $V_{t+1}^M\left(a+1, e_i, s_{i,a+1,t+1}\right)$, when computing $V_t\left(a, e_i, s_{i,a,t}\right)$ and $V_t^M\left(a, e_i, s_{i,a,t}\right)$. As a result, value function iteration and recursive age-based iteration are not needed. Instead, the Bellman equations are directly solved on the joint state variable grid using the full transition matrix constructed in Section IA2.1. The technology adoption policy, m_t, and expected value of entry are determined after solving for $V_t\left(0, e_i, s_{i,0,t}\right)$ and $V_t^M\left(0, e_i, s_{i,0,t}\right)$ for entrants. The labor policy functions, n_t, n_t^M, and exit policy functions, x_t, x_t^M, are determined while computing V_t and V_t^M. The equilibrium price p_t is determined using `Broyden` function to satisfy the free entry condition (3) or (8) as stationary equilibrium. This backward iteration continues until the initial steady state is reached.

The initial firm distribution $\bar{\mu}_0$, $\bar{\mu}_0^M$ is determined in the initial steady state. Given the solved equilibrium price sequences, $\{p_t\}_{t=1}^{T-1}$, and policy function sequences, $\left\{n_t, x_t, n_t^M, x_t^M, m_t\right\}_{t=1}^{T-1}$, we then solve forward for $\left\{\bar{\mu}_t, \bar{\mu}_t^M, \bar{M}_t\right\}_{t=1}^{T-1}$ using the product market clearing condition (5) or (11) and the laws of motion (9) and (10). This forward iteration continues until the terminal steady state is reached.

The initial steady state is set as 1859 and the terminal steady state is 2109 for all models, with a transition period of 250 years. This time frame captures key historical points, 1917, 1955, and 2018.

IA2.4 Model Calibration

To calibrate the models in Sections 5.1, 5.2, and 5.3, our objective function takes 13 parameters in each row of Table 2, Panel A, as inputs, and outputs the weighted mean squared error of the 13 moments in Panel B. We rescale the median age by a factor of $\frac{1}{100}$ to keep all moments numerically comparable. All moments are equally weighted except for revenue growth rates by age, which have a relatively small magnitude compared to exit rate moments. To account for this, we assign revenue growth a weight five times higher. The step tolerance and optimality tolerance are set as 10^{-6}.

To efficiently select the initial guess in optimization, we first construct a Sobol sequence with 2^{13} points over the state space to ensure better coverage and avoid clustering issues. We then identify points with smaller residuals and form candidate balls around them. Initial guesses are drawn from Sobol sequences and uniform samples within these candidate balls. Optimization is performed using MATLAB's `fmincon`. If firm number growth surpasses 50% between 1955 and 2018, or if the firm exit rate drops below 0.01 or revenue growth turns negative, we apply a penalty to drive the parameters out of this region. We exclude any calibration where the startup rate is zero at any point along the transition path.

The calibration for in Section 6 follows the same procedure but is relatively less computationally intensive. Our objective function takes nine or ten parameters in each row of Table 3, Panel A (depending on whether the learning rate is calibrated) as inputs, and outputs the weighted mean squared error of seven or eight moments in Panel B. The median age is again rescaled by a factor of $\frac{1}{100}$ so that all moments are of similar magnitude, and all moments are equally weighted. The step tolerance and optimality tolerance are set as 10^{-6}.

IA2.5 Model Simulation

This section describes the method used to simulate the model, in order to examine the overlap among top firms in 1955 and 2018. The same method applies to the models described in Sections 5.1, 5.2, and 5.3. The solved policy functions, n_t, x_t, for traditional technology are used in simulating all model variants. The counterparts associated with modern technology, n_t^M, x_t^M, are used only in Section 5.1.

We simulate the nonstationary transition path and track firm evolution using unique firm IDs. Firms are initially drawn from the stationary distribution in the initial steady state solved in Section IA2.2, with IDs assigned sequentially starting from 1. In each period, incumbent states evolve through a Markov chain simulation, where the transition matrix are constructed over the joint state variable grid as described in Section IA2.1. Incumbents in the next period face exogenous exit and make endogenous exit decisions based on their exit policy function x_t, x_t^M solved in Section IA2.3 after observing their next-period states.

New entrants in the next period are drawn with assigned productivity states and make technology choice according to the entry probability vector constructed in Section IA2.1. After observing the productivity draw and selecting the technology type, entrants decide whether to proceed with production based on their exit policy functions x_t, x_t^M, solved in Section IA2.3. Those that remain in the market are assigned consecutive IDs, continuing from the highest existing firm ID to maintain uniqueness. To ensure market clearing, new entrants must fill the gap between total demand and the output produced by surviving incumbents. Output is computed using labor policy functions n_t, n_t^M solved in IA2.3. Demand grows according to market growth rate, η_t. We draw entrants iteratively until equilibrium is reached.

This process is repeated for 160 years, from 1859 to 2018, with a burn-in period of 500 years to establish a stable firm distribution in the initial steady state. We simulate the nonstationary equilibrium 500 times using different random seeds across multiple CPU cores and average the results across simulations.