The Global Change in the Corporate Production Function*

Erica X.N. Li† Guoliang Ma‡ Dan Su§

June 17, 2025

Abstract

The assumption of a concave production function has long been a cornerstone in economics literature. However, using Bayesian Markov Chain Monte Carlo (MCMC) to estimate structural breaks, we find that since the 1980s, corporate production functions have transitioned to a sigmoidal (convex-concave) shape, with the convex region becoming increasingly dominant over time. This structural shift, observed across various countries and industries, quantitatively explains the growing share of firms with negative net earnings. Lastly, we explore the implications of this transformation for corporate market power.

JEL codes: D22; G12; G30; L20; O33

Keywords: corporate production function; Bayesian MCMC; return to scale; negative earnings; market power

*First Version: November 2023. For helpful comments and suggestions, we thank Filippo Biondi, Giulio Gottardo, Yu Li, Yang Liu (discussant), Gordon Phillips, and other participants at Chinese University of Hong Kong, European Association for Research in Industrial Economics Annual Conference, European Economic Association Annual Meeting, European Winter Meeting of the Econometric Society, Five-Star Workshop in Finance, Five-Star Asia Pacific Workshop in Finance, FIW-Research Conference International Economics, Hong Kong Baptist University, and University of Macau. Of course, all errors are our own.

†Department of Finance, Cheung Kong Graduate School of Business. Email address: xnli@ckgsb.edu.cn

‡The Paula and Gregory Chow Institute for Studies in Economics, Xiamen University. Email address: glma@xmu.edu.cn

§Department of Finance, Cheung Kong Graduate School of Business. Email address: dansu@ckgsb.edu.cn

1 Introduction

The concept of returns to scale has long been central to the study of production and firm dynamics. Since the foundational work of Charles W. Cobb and Paul H. Douglas (1928), the assumption of a (quasi-)concave production function—such as the Cobb-Douglas or Constant Elasticity of Substitution (CES) forms—has been a cornerstone in macroeconomic models, shaping our understanding of economic growth, resource allocation, and business cycles. However, our study documents a profound shift in the shape of corporate production function since the 1980s, as it evolves toward a sigmoidal (convex-concave) structure. This shift, marked by the increasing prominence of the convexity component, challenges the traditional concave production function assumption and suggests a fundamental reconfiguration in how firms *produce*, *invest*, and *manage earnings dynamics*. Our analysis reveals that this shift is not confined to a specific industry or country but is instead a widespread phenomenon observed across multiple countries and industries, which suggests a global transformation in the production landscape.

We begin by documenting a key motivating fact: the share of unprofitable firms has risen significantly over the past several decades. Using data on publicly traded firms in the U.S., we show that the proportion of firms with negative net earnings increased from 18.3% in 1980 to 54.4% in 2019. This trend is not confined to a single industry in the U.S.; similar upward patterns are observed across industries and economies worldwide. We interpret this fact through the lens of Q-theory, and show that if the production function has become sufficiently convex, firms could have negative earnings – an equilibrium outcome that would not occur under a strictly concave production function. Based on this, we hypothesize that the long-term increase in unprofitable firms is driven by structural changes in the shape of the corporate production function.

Our analysis examines the long-term evolution of the corporate production function. Our baseline study utilizes firm-level data from publicly traded U.S. companies and extends the investigation to multiple advanced and emerging economies. More specifically, employing a Bayesian Markov Chain Monte Carlo (MCMC) estimation method with minimal structural assumptions on the production function, we uncover two key findings. First, we provide empirical evidence that the corporate production function consists of two distinct regions. Before transitioning into the usual concave, decreasing returns-to-scale phase, firms experience an initial convex region characterized by increasing returns to scale. In our baseline analysis using U.S. data, the estimated average degree of returns to scale in this convex region is 1.11, significantly above 1.0 at the 5% significance level for most periods, while the estimate for the concave region is 0.99 with borderline significance. In other words, the corporate production function follows a convex-concave shape rather than a purely concave one. Second, we identify a significant transformation in the production function's shape since 1980. In our baseline analysis, the estimated convexity-concavity threshold was 5.02 thousand dollars in 1980; by 2021, it had surged to 1.31 million dollars, marking a remarkable 261-fold increase. Over the same period, the average returns-to-scale rose from 0.97 to 1.07, indicating that modern firms now operate under an extended phase of increasing returns to scale, requiring a substantially larger operational scope before diminishing returns set in. This shift remains robust across alternative functional specifications and is consistently observed across various industries and economies. Notably, the *Healthcare*, *Manufacturing*, and *Consumer Nondurables* sectors exhibit particularly strong patterns, as do economies such as *China*, *India*, and *Australia*.

Building on this empirical evidence, we conduct a quantitative exercise to explore the broader economic implications of this evolving production function. By using an otherwise standard firm-dynamics model a la Hugo A. Hopenhayn (1992), we simulate firm dynamics under a convex-concave production function with capital adjustment costs, we illustrate how the shift toward convexity impacts firm behavior and market outcomes. Our results suggest that the rising share of firms with negative net earnings can be attributed to increased capital investment among highly productive firms with small capital stocks striving to reach profitable scales. These firms, aiming for growth, often incur short-term losses due to the substantial upfront costs of capital accumulation in a more convex production environment. This quantitative exercise highlights how the transformation in production function shape contributes to a higher prevalence of unprofitable firms, a trend consistent with our empirical findings.

Finally, we explore how changes in the shape of the production function influence markup estimates, with a particular focus on disentangling the effects of technological change from those of rising market power. Building on the findings of Jan De Loecker, Jan Eeckhout and Gabriel Unger (2020), who document a sharp increase in aggregate markups—from 21% above marginal cost to 61% in recent years—we argue that much of this observed rise may be attributed to technological advancements rather than to a pure increase in corporate market power. Their methodology, which assumes constant returns to scale and does not fully account for shifts in production technology, may conflate technological progress with changes in firm-level pricing behavior. This concern has been echoed by James Traina (2021), who show that after adjusting for marketing and management expenses, public firm markups increased only modestly and remained within historical variation. Furthermore, Steve Bond, Arshia Hashemi, Greg Kaplan and Piotr Zoch (2021) argue that when only revenue data are available, output elasticities cannot be non-parametrically identified under market power, complicating the interpretation of rising markups. Additionally, Steve Bond and Giulio Gottardo (2024) demonstrate that the observed markup increase reported by De Loecker, Eeckhout and Unger (2020) is heavily influenced by the use of arithmetic averages;

when the harmonic mean is employed, aggregate markups appear relatively stable over time. In our model-based simulations, we show that even when all observed changes in the convexity of the revenue production function stem solely from technological improvements—absent any variation in market power—the estimated average markup still rises over time. Notably, our model highlights that although technological factors explain a significant portion of the aggregate markup increase, they cannot fully account for the empirical patterns observed in the data. Specifically, empirical evidence indicates that the upward trend in markups began earlier, around the 1980s, and reached higher levels—up to 1.60—compared to the 1.40 ceiling suggested by our simulations. Moreover, De Loecker, Eeckhout and Unger (2020) document that the sharpest markup increases occur at the upper tail of the distribution, with the 90th percentile reaching levels as high as 2.5, contrasting with the uniform percentile trends implied by our technology-driven model. These discrepancies, along with evidence of early-onset markup increases and sharper distributional shifts, underscore the role of additional drivers such as regulatory changes, market concentration, and strategic firm behavior. Nonetheless, our results provide robust evidence that technological forces—manifested in increased returns to scale and reduced variable costs—are a central driver of the long-run rise in markups, challenging the notion that market concentration alone is responsible for this trend.

Related literature Our paper is closely related to four branches of literature. First, our work contributes to the growing literature on the evolving characteristics of firms in the 21st century. De Loecker, Eeckhout and Unger (2020) document a substantial increase in market power among U.S. public firms since 1980, while Gerard Hoberg and Gordon Phillips (2021) show that firms have significantly expanded their scope and scale of operations over the past 30 years. Additionally, David Autor, David Dorn, Lawrence F. Katz, Christina Patterson and John Van Reenen (2020) highlight the rising dominance of superstar firms that lead their markets, and both Callum Jones and Thomas Philippon (2016) and German Gutierrez and Thomas Philippon (2017) provide evidence of declining competition and investment among U.S. firms. Our work complements this body of research by documenting another key trend of rising unprofitable companies in the changing landscape of modern firms.

Second, our paper contributes to the ongoing debate on rising corporate market power. De Loecker, Eeckhout and Unger (2020) document a sharp increase in aggregate markups—from 21% above marginal cost to 61% in recent years. However, this trend has been challenged by several studies. For example, Traina (2021) argue that once marketing and management expenses are properly accounted for, the rise in public firm markups is modest and within historical bounds. Bond et al. (2021) highlight that when only revenue data are available, output elasticities are not non-parametrically identified under market power. Bond and Gottardo (2024) further show that the increase reported in De Loecker, Eeckhout and Unger (2020) is largely driven by the use of arithmetic averages; when the harmonic mean is used, aggregate markups appear more stable. Our contribution is to demonstrate that changes in technology—rather than increases in market power per se—can also generate the observed rise in markups.

Third, our paper aligns with the recent literature on increasing returns-to-scale. Empirical and theoretical studies have documented substantial variation in returns to scale across sectors and firm sizes. For instance, Chengyu Gao and Matthias Kehrig (2021) report an average return to scale of 0.96 in U.S. manufacturing, with variability across four-digit industries. Similarly, Dimitrije Ruzic and Kwan Yeung Ho (2019) observe a decline in returns to scale in U.S. manufacturing, suggesting structural shifts in production technology or market conditions. European studies, such as Danial Lashkari, Max Bauer and Vianney Boussard (2019) on France, also reveal significant heterogeneity in returns to scale, particularly across firms with different levels of IT investment, which affect the elasticity of output relative to inputs. In addition, Joel Kariel, Anthony Savagar and Joao Mainente (2022) offers a comprehensive analysis of returns to scale across sectors in the UK, finding slightly above-one returns-to-scale estimates, indicating increasing returns-to-scale on average. Our contribution is to provide complementary global evidence by using Bayesian MCMC approach.

Finally, our work contributes to the production function estimation literature. Previous studies have focused on estimating firm-level TFP by developing various methods to address simultaneity and selection biases (e.g., G Steven Olley and Ariel Pakes, 1996; James Levinsohn and Amil Petrin, 2003; Jeffrey M Wooldridge, 2009; Daniel A Ackerberg, Kevin Caves and Garth Frazer, 2015). In contrast to earlier work, our Bayesian MCMC approach offers greater flexibility, enabling us to capture time-series changes in production functions dynamically. Moreover, our firm-level analysis complements prior country-level estimates of aggregate production functions (e.g., Robert Solow, 1957; Paul Samuelson, 1979; Robert E. Hall and Charles I. Jones, 1999), providing new insights into micro-level production dynamics. In addition, our paper connects to the literature on convex-concave production functions. This structure is especially relevant in development economics, where economies transition from low levels of development with increasing returns to higher levels marked by diminishing returns. Foundational theoretical contributions include Zvi Griliches (1957), William Ginsberg (1974), and A. K. Skiba (1978), among others. Subsequent studies have applied convex-concave production functions to explain poverty traps, where economies are stuck at low income levels (e.g., Costas Azariadis and Allan Drazen, 1990; Philippe Askenazy and Cuong Le Van, 1999; Ken-Ichi Akao, Takashi Kamihigashi and Kazuo Nishimura, 2011). In these models, whether an economy converges to a high or low steady-state equilibrium depends on its initial capital per capita. While much of this research is theoretical and focuses on country-level analysis, our study provides the first firm-level evidence of the increasing significance of the initial convex phase in production functions over time.

Layout The rest of our paper is organized as follows. Section 2 outlines the data sources and variables utilized in our analysis. Section 3 introduces the motivating fact and our interpretation through the lens of Q-theory. Section 4 details our Bayesian methodology for empirical analysis and summarizes our key findings regarding shifts in the production function of public firms globally. In Section 5, we employ a standard firm dynamics model to assess the quantitative performance of our story and explore its implications on corporate market power. Finally, Section 6 provides concluding remarks.

2 Data and Variable Construction

2.1 U.S. Evidence

The dataset used in our empirical analysis for the U.S. primarily comes from *Compustat*, which offers comprehensive balance sheet data for publicly listed companies. We focus on firms with a foreign incorporation code of "USA" and exclude financial institutions (SIC 6000-6999) and regulated utilities (SIC 4900-4999) in our baseline analysis. Moreover, we exclude firms with missing or negative values for total assets or sales.

All variable definitions adhere to standard practices in corporate finance literature. Specifically, a firm's output is defined as net sales or turnover (*Compustat* data item *SALE*), and firm size is measured as the natural logarithm of total assets (*Compustat* data item *AT*). Net earnings are captured by *Compustat* data item *NI*, which reflects a firm's income or loss after accounting for all revenues, gains, expenses, and losses. Gross profit (*Compustat* data item *GP*), in contrast, only subtracts the cost of goods sold (*Compustat* data item *COGS*) from total revenue (*Compustat* data item *REVT*). Firm age is determined based on the first year the firm appears in *Compustat*.

To account for the growing importance of intangible assets, a firm's total capital stock is defined as the sum of tangible capital (*Compustat* data item *PPENT*) and intangible capital, measured as in Ryan H. Peters and Lucian A. Taylor (2017). Thus, investment includes both tangible and intangible capital. Specifically, total investment I is calculated as $I = K^{tangible} - K^{tangible}_{-1} + DP - AM + XRD + 0.3 \times SG\&A$, where the first two terms represent the growth of physical capital stock, DP is depreciation and amortization (*Compustat* data item *DP*), AM is amortization (*Compustat* data item *AM*), XRD denotes R&D costs (*Compustat* data item *XRD*), and $SG\&A$ represents net selling, general, and administrative expenses (net SG&A). This net SG&A is calculated by excluding data items *XRD* and *RDIP* (In Process R&D Expenses) from data item *XSGA* in *Compustat*, as *XSGA* in *Compustat* often includes R&D costs.[1] We set organization capital investment to 30% of SG&A expenditures, following Andrea L. Eisfeldt and Dimitris Papanikolaou (2013) and Peters and Taylor (2017).

For industry classification, we use the Fama-French ten-industry approach.[2] For industry-level analysis, financial institutions (SIC 6000-6999) and regulated utilities (SIC 4900-4999) are reintegrated into the dataset. Additionally, all initial public offerings (IPO)-related information is sourced from Jay Ritter's publicly accessible data.[3]

2.2 Global Evidence

For our global analysis, all country-level information is obtained from the Penn World Table (PWT). We use output-side constant-price real GDP divided by total population to capture cross-country differences in real GDP per capita and use the price level of capital stock to account for changes in the cost of capital formation due to technical advancements or inflation.

For international firms, data is drawn from the *Compustat Global* dataset, where we apply similar data-cleaning criteria as used for U.S. firms. The *Compustat Global* dataset provides extensive firm-level balance sheet data for publicly listed companies in over 80 countries, collectively covering more than 90% of global market capitalization. Given that our Bayesian method requires substantial data, we set a minimum threshold for firm-year observations. Specifically, a country is included in our sample if it has at least 10,000 firm-year observations. Based on this criterion, the countries included in our international analysis are *Australia*, *Canada*, *China (including Hong Kong and Taiwan)*, *France*, *Germany*, *India*, *Japan*, *Korea*, *Sweden*, *Thailand*, and *UK*.

All variable definitions remain consistent with those used in our analysis of the U.S. dataset. The only distinction here is that we need to construct the intangible capital stock independently. Following Peters and Taylor (2017), we measure intangible capital stock using the perpetual inventory method. To begin with, we estimate the initial intangible capital stock with the following

[1]If *XSGA* is missing, we set *SG&A* to zero. Additionally, if *XRD* exceeds *XSGA*, we set *SG&A* equal to *XSGA*.

[2]The definitions for these industries are as follows: *Consumer Nondurables* (SIC 0100-0999, 2000-2399, 2700-2749, 2770-2799, 3100-3199, 3940-3989); *Consumer Durables* (SIC 2500-2519, 2590-2599, 3630-3659, 3710-3711, 3714-3714, 3716-3716, 3750-3751, 3792-3792, 3900-3939, 3990-3999); *Manufacturing* (SIC 2520-2589, 2600-2699, 2750-2769, 2800-2829, 2840-2899, 3000-3099, 3200-3569, 3580-3621, 3623-3629, 3700-3709, 3712-3713, 3715-3715, 3717-3749, 3752-3791, 3793-3799, 3860-3899); *Oil, Gas, and Coal Extraction and Products* or *Energy* (SIC 1200-1399, 2900-2999); *Business Equipment* (SIC 3570-3579, 3622-3622, 3660-3692, 3694-3699, 3810-3839, 7370-7372, 7373-7373, 7374-7374, 7375-7375, 7376-7376, 7377-7377, 7378-7378, 7379-7379, 7391-7391, 8730-8734); *Telephone and Television Transmission* (SIC 4800-4899); *Wholesale, Retail, and Some Services* (SIC 5000-5999, 7200-7299, 7600-7699); *Healthcare, Medical Equipment, and Drugs* (SIC 2830-2839, 3693-3693, 3840-3859, 8000-8099); *Utilities* (SIC 4900-4949); and *Others*.

[3]https://site.warrington.ufl.edu/ritter/ipo-data/

equation:

$$K_0^{Intangible} = \frac{I_0^{Intangible}}{g + \delta^{Intangible} - \pi} \tag{1}$$

In the equation above, $I_0^{Intangible}$ represents the firm's investment in organizational capital in the first sample year. Consistent with our previous approach, we set organizational capital investment to 30% of $SG\&A$. The term g denotes the industry-level average growth rate of SG&A investments, using the first 2-digits of the NAICS industry code to classify the industry within each country. The term π represents the growth rate of the capital stock price, accounting for changes in the real cost of capital investment. For the intangible depreciation rate, $\delta^{Intangible}$, we use 20% following Eisfeldt and Papanikolaou (2013). Once the initial intangible capital is obtained, we iterate forward using the depreciation rate, SG&A expenses, and investment price index with the following equation:

$$K_{t+1}^{Intangible} = (1 - \delta^{Intangible}) K_t^{Intangible} \pi_{t+1} + I_{t+1}^{Intangible} \tag{2}$$

3 Facts and Interpretation

3.1 Facts

To start, we document the increasing prevalence of unprofitable firms as a motivation fact. Figure 1 presents our baseline analysis of the time-series trend in the proportion of firms reporting negative net incomes. For each year, we calculate the fraction of firms with negative net incomes by dividing the number of such firms by the total number of firms in our sample. We use two distinct measures for this calculation: one weighted by each firm's industry output share and an unweighted measure.

[Figure 1 here]

As shown in Graph (a) of Figure 1, there is a persistent increase in the share of firms with negative earnings in both measures. For the unweighted indicator (i.e., the blue solid line), 18.3% of firms reported negative net incomes in 1980, a proportion that climbed substantially to 54.4% by 2019. Similarly, the weighted measure (i.e., the orange solid line) reveals an increase from 14.8% in 1980 to 37.4% in 2019. Although there was a significant drop around the year 2000, this upward trend in unprofitable firms has strengthened in recent years. With these two simple measures, our analysis reveals a long-term rise in the proportion of unprofitable public firms in the U.S., which signals a fundamental shift in the earnings landscape. This trend suggests that (seemingly) unprofitability is no longer a rare or temporary phenomenon for public companies but rather a persistent feature that warrants further exploration.

robustness checks We perform several robustness checks to confirm the reliability of our findings. First, we show that the observed upward trend in the share of unprofitable firms is not limited to a specific industry. Figure A1 in the appendix presents the proportion of firms with negative net incomes across the Fama-French ten industries. As illustrated, the fraction of unprofitable firms has steadily increased in most industries, with particularly notable rises in the *Healthcare* sector (from 20% to over 80%) and the *Business Equipment* sector (from just below 20% to 60%). In contrast, this trend is less pronounced in the *Utilities* sector, and it fluctuates considerably over time in sectors such as *Energy* and *Other*. Some sectors, including *Manufacturing* and *Telephone and Television Transmission*, saw peaks around the year 2000, followed by gradual declines. Despite these variations across industries, the general upward trend in the share of unprofitable firms remains consistent and is not confined to any single sector.[4] This broad-based increase underscores the pervasiveness of unprofitability across various segments of the economy and suggests a structural shift affecting multiple industries.

Next, we investigate whether the rising trend in unprofitable firms is influenced by an increasing proportion of younger firms within the *Compustat* dataset. The presence of more young public firms today, which generally have lower net earnings, could potentially drive the observed increase in unprofitable firms. To address this concern, we examine two age-related indicators in Figure A2 in the appendix. The first indicator, shown by the yellow line, is the average firm age. This measure reveals an upward trend over time, indicating that public firms have been aging on average, with a larger proportion of mature firms in recent years. This finding suggests that the public firms are becoming older, not younger. The second indicator, represented by the green line, is a proxy for the proportion of young firms, defined as the share of firms that are five years old or younger. Although this indicator exhibits some fluctuations, it does not show a clear upward trend over time. These findings imply that the observed increase in the share of unprofitable firms is not driven by changes in the age distribution of public firms. Instead, it appears that the trend towards unprofitability is widespread and persists even as the public firm population becomes older, underscoring that this phenomenon is not merely an effect of having more young, early-stage firms in the dataset.

Lastly, we investigate whether the observed pattern of rising unprofitability is concentrated within specific stock exchanges, as listing requirements—particularly financial criteria—vary across

[4] For further context, Table A2 in the appendix provides a list of the top 50 companies by market capitalization with negative net earnings in 2019 across multiple industries, including *Agriculture, Manufacturing, Retail Trade, and Services*.

exchanges. Figure A3 in the appendix illustrates this analysis by displaying the proportion of firms with negative net earnings on the New York Stock Exchange (NYSE), the National Association of Securities Dealers Automated Quotations (NASDAQ), and other U.S. exchanges. The red line represents the NYSE, where the share of unprofitable firms grew from 10.5% in 1970 to 31.4% in 2019. In contrast, the green line for NASDAQ shows a more pronounced increase, with the proportion of unprofitable firms rising from 15.5% to 63.7% over the same period. These differences highlight some heterogeneity across exchanges; firms listed on NASDAQ, for example, have generally exhibited a higher incidence of unprofitability than those on the NYSE. Despite these variations, our overall conclusion regarding the secular rise in the fraction of unprofitable firms remains robust across different stock exchanges. This consistent trend suggests that the increase in unprofitability is a broad-based phenomenon affecting firms across multiple exchanges, regardless of varying listing standards.

gross v.s. net Interestingly, this upward trend is much less pronounced when we focus on the proportion of firms with negative gross profits. As shown by the two dotted green and purple lines in Graph (a) of Figure 1, although the percentage of firms with negative gross profits has risen over the past decades, their overall economic impact remains relatively limited. Specifically, under the unweighted measure, the share of firms with negative gross profits increased from 1.7% in 1980 to 10.2% in 2019. For the weighted measure, this figure rose from 1.3% to 3.3% over the same period. Therefore, despite the challenges many firms face, reflected in negative or unusually low net earnings, most public firms still report positive gross profits.

The divergence between net earnings and gross profits is crucial for understanding the underlying mechanisms at play. As we elaborate in the following section, in theory, this gap between profitability measures could result from a fundamental shift in the shape of the production function. Intuitively, when a firm reports positive gross profits but negative net earnings, it indicates that while its core operations remain profitable, significant resources are being allocated toward scaling the business. Under a concave production function, this pattern would be rare, as marginal returns on investment diminish as the firm expands. However, in the presence of a convex production function, profitability increases with firm size, implying that a firm that is profitable today could achieve even greater profitability as it grows. As a result, firms in this environment may prioritize aggressive investment and expansion, even at the expense of current earnings.[5]

[5]Moreover, the widening gap between gross profits and net earnings provides insight into the ongoing debate on firm-level markup measurement. While De Loecker, Eeckhout and Unger (2020) document a significant rise in corporate markups over recent decades, other studies, such as Traina (2021), present conflicting findings. This discrepancy largely stems from differences in the measurement of input costs. Traina (2021) incorporate operating expenses, while De Loecker, Eeckhout and Unger (2020) focus on the cost of goods sold. Notably, operating expenses include marketing

evidence from IPO We extend our analysis by investigating trends in IPO performance. Graph (b) of Figure 1 displays the fraction of U.S. firms with negative net earnings at the time of their IPOs. Following standard practice in the IPO literature, corporate earnings are measured based on the most recent twelve months before going public. Then the fraction is simply computed by dividing the number of IPO firms reporting net losses by the total number of firms that went public in a given year. The solid blue line in Graph (b) illustrates this trend over time, which reveals a clear upward trajectory: while only 24% of firms had negative net earnings at IPO in 1980, this figure has surged to 77% by 2019.

Importantly, this rise in unprofitable IPOs is not solely attributable to the expansion of the IT sector. The red dashed line in Graph (b) of Figure 1 tracks the share of IT-related IPOs over time. Before 2000, the increase in unprofitable IPOs appeared to be largely driven by the growing presence of IT firms. However, after 2000, this relationship weakens significantly. While the proportion of IPO firms with negative earnings has continued to climb, the share of IT-related IPOs has remained relatively stable. One possible explanation is the rise of high-growth, non-traditional firms outside the IT sector—such as Tesla and Peloton—that have also pursued IPOs despite reporting losses. This pattern aligns with our earlier findings in Figure A1, which document a sustained increase in the share of unprofitable firms across various industries. Collectively, these results underscore the broad-based nature of our documented trend, suggesting that IPO unprofitability has become a widespread phenomenon across many sectors, rather than being concentrated within the technology industry.

global evidence There is a possibility that data on U.S. publicly traded firms may be subject to selection bias, suggesting that the observed patterns could be specific to the U.S. To address this concern, we extend our previous analysis by using a global firm-level dataset. The main findings are presented in Graph (c) of Figure 1, which displays the time-series trends in the fraction of firms reporting negative net earnings and negative gross profits globally. As in our previous analysis, we provide two versions of each measure: one weighted by the relative industry importance of the firm and an unweighted series. The solid lines in Graph (c) reveal a global increase in the share of firms with negative net earnings across both measures. Specifically, in the unweighted series, the proportion of firms with negative net income rose from 2.7% in 1987 to 29.6% in 2019, while the weighted measure similarly increased from 1.1% to 26.4% over the same period. In contrast, the rise in firms with negative gross profits is much less pronounced, as shown by the dashed lines

and management costs in addition to production costs. We argue that firms increasingly treat sales and marketing expenditures as long-term investments in customer acquisition to strengthen future market power. Consequently, these expenses should be excluded from current markup calculations to provide a more accurate assessment of market power dynamics and profitability potential. We put more discussions in Section 5.3.

in the same graph. For instance, the unweighted measure for firms with negative gross profits increased from 0.8% in 1987 to 5.5% in 2019, while the weighted measure rose from 0.2% to 4.3%. These findings confirm that the trend observed among U.S. firms is not unique but rather part of a broader, global phenomenon.

Another noteworthy finding is the cross-sectional relationship between a country's real GDP per capita and the prevalence of firms reporting negative net earnings. Graph (d) of Figure 1 presents a binned scatter plot of log real GDP per capita against the share of unprofitable firms by country. The fitted linear trend is represented by the blue dashed line. The data reveal a significant positive relationship, which indicates that firms in wealthier countries are more likely to report earnings losses. This cross-country pattern suggests that the rise in unprofitable firms is not solely driven by low institutional quality or weak corporate governance. Instead, both supply- and demand-side factors appear to play a crucial role. On the supply side, wealthier economies tend to have a higher concentration of high-tech and e-commerce firms, which heavily rely on intangible assets and operate under increasing returns to scale. These firms often incur substantial upfront costs to develop user networks and infrastructure, leading to short-term earnings losses as they strive to establish market dominance. On the demand side, emerging economies generally have less developed financial markets and stricter IPO regulations, resulting in higher listing standards on financial terms. Consequently, firms with negative net earnings are less likely to receive IPO approval in these countries. Taken together, these supply- and demand-side factors help explain the observed positive cross-country relationship in the prevalence of unprofitable firms.

3.2 A Q-theory Interpretation

3.2.1 Some clarifications on degree of returns-to-scale and investment

We now turn to the theoretical explanation of the underlying drivers behind the rise in unprofitable firms. To clarify, we begin by defining our concepts of returns to scale and capital investment. Consider a firm that utilizes capital K_t and labor L_t to produce a non-storable output $\tilde{Y}_t$, which is sold at a market price of $\tilde{P}_t$ at time t. The firm's *quantity* production function is specified as

$$\tilde{Y}_t = \tilde{A}_t \left(K_t^{\gamma_t} L_t^{1-\gamma_t} \right)^{s_t}$$

where $\tilde{A}_t$ represents productivity at time t. Here, the parameter $0 < s_t \leq 1$ reflects the degree of returns to scale in *quantity* output production, while $0 < \gamma_t < 1$ denotes the relative factor shares of labor and capital.

Changes in returns to scale, denoted by s_t, can originate from variations in either fixed costs or marginal costs or both. The intuition behind this argument is that as output expands, fixed costs are spread over more units and marginal costs decline due to scale efficiencies, both of which lower average cost. Consequently, total cost increases less than proportionally with output, satisfying the condition for increasing returns to scale. Details are provided in Appendix B.1.

The inverse demand function for the firm's output is given by

$$\tilde{P}_t = \left(\frac{\tilde{Y}_t}{\tilde{H}_t}\right)^{-\frac{1}{\varepsilon_t}}$$

where $\tilde{H}_t > 0$ positions the demand curve, and $\varepsilon_t \geq 1$ is the price elasticity of demand.

As shown in Appendix C.1, given the wage rate w_t, by optimizing labor choice L_t, we can express maximized sales incomes Y_t as:

$$Y_t \equiv \tilde{P}_t \tilde{Y}_t = \tilde{Z}_t^{1-\alpha_t} K_t^{\alpha_t}$$

where $\tilde{Z}_t$ is a function of $\tilde{A}_t$, $\tilde{H}_t$, and w_t. The *returns to scale* α_t for sales Y_t with respect to capital K_t is defined as:

$$\alpha_t \equiv \frac{\gamma_t s_t \left(1 - \frac{1}{\varepsilon_t}\right)}{1 - (1 - \gamma_t)\, s_t \left(1 - \frac{1}{\varepsilon_t}\right)} > 0 \tag{3}$$

where α_t is our definition of return-to-scale throughout our analysis. Later, we provide a detailed explanation of how α_t is measured in the data and discuss its structural changes over time.

Our benchmark scenario considers a fully competitive firm ($\varepsilon_t = \infty$) operating under constant returns to scale in production ($s_t = 1$) with stable factor shares ($\gamma_t = \bar{\gamma}$). In this case, $\alpha_t = 1$, signifying a constant returns-to-scale production function for sales income or output measured in dollars. We use α_t as the return-to-scale parameter in our analysis because the typical cross-country dataset does not allow us to separately observe quantity and price components. Thus, if we observe changes in α_t, they may stem from multiple factors. Our current approach does not pinpoint the exact source, so variations in our estimated α_t could reflect shifts in corporate market power ε_t (De Loecker, Eeckhout and Unger, 2020), adjustments in the *quantity* returns-to-scale parameter s_t, or changes in factor shares γ_t (Loukas Karabarbounis and Brent Neiman, 2014). For example, if a firm has some degree of monopoly power ($\varepsilon_t < \infty$) or operates under decreasing returns to scale in quantity ($s_t < 1$), then $\alpha_t < 1$, making sales a concave function of capital and generating positive economic rents. Conversely, if the firm experiences increasing returns to scale in quantity production ($s_t > 1$) and grows large enough, we may observe $\alpha_t > 1$

in the data, with sales then exhibiting a convex relationship with capital stock. In other words, α_t could be larger than 1 even in the presence of monopoly power ($\varepsilon_t < \infty$). We put more discussions on the implications of our approach on corporate market power in Section 5.3.

In this analysis, capital comprises both intangible and tangible assets. Thus, the term "investment" here may include physical capital investment as well as expenditures on customer capital, organizational capital, and other intangible assets. This approach is consistent with our empirical method, where we do not differentiate between tangible and intangible capital stocks and their corresponding investments. Additionally, we assume capital is the firm's only inflexible input, while all other inputs, such as labor and materials, are frictionless. Consequently, it does not affect our results on estimating α_t whether we use sales (i.e., Y_t) or value-added ($Y_t - w_t L_t$) as our output measure.

3.2.2 The economics behind negative earnings

Now we turn to a Q-theory interpretation on why firms could make negative earnings and under what conditions. Consider a infinite-horizon continuous-time economy with exogenously-determined interest rate $\{r_t\}_{t=0}^{\infty}$. Let the sales of a single firm be given by $Y_t = F(K_t)$ and the increase per time unit in the firm's capital stock is given by $\dot{K}_t = I_t - \delta K_t$ with $K_0 > 0$. I is gross fixed capital investment per time unit and $\delta > 0$ is the rate of depreciation of capital. There are adjustment costs associated with investment, and they are denoted by $G(I, K)$. The installation cost function, G, is a C^2 function satisfying the usual assumptions, i.e., $G(0, K) = 0$, $G_I(0, K) = 0$, $G_{II}(I, K) > 0$, $G_K(I, K) \leq 0$, for all pairs (I, K) with $I \geq 0$ and $K \geq 0$. At the cost of some minor and uninteresting loss of generality, we further assume that G is a (jointly) convex function of (I, K) and homogeneous of degree 1 in its respective arguments. It means that $G_{KK} \geq 0$, $G_{II}G_{KK} - (G_{IK})^2 \geq 0$, and $G(I, K) = G_I(I, K)I + G_K(I, K)K$, for all (I, K). Examples of G functions satisfying these assumptions are the widely-used quadratic adjustment cost function such as $G(I, K) = \frac{1}{2}\beta\frac{I^2}{K}$, where $\beta > 0$ denotes the degree of investment inflexibility.

Let the net earnings at time t be denoted π_t, and the gross profit at time t be R_t, then we have

$$\begin{aligned} \pi_t &\equiv F(K_t) - G(I_t, K_t) - I_t \\ R_t &\equiv F(K_t) \end{aligned}$$

In this way, the decision problem, as seen from time 0, is the following: given the expected evolution of interest rates, $\{r_t\}_{t=0}^{\infty}$, choose an investment plan $\{I_t\}_{t=0}^{\infty}$ so as to maximize the firm's market value, i.e., the discounted value of the future stream of expected net earnings:

$$\max_{\{I_t\}_{t=0}^{\infty}} V_0 = \int_0^{\infty} \pi_t e^{-\int_0^t r_\tau d\tau} dt \tag{4}$$

subject to the constraints mentioned previously.

With this model setup, we can have the following lemma on the relationship between the shape of production function $F(K_t)$ and the sign of net earnings π_t.

Lemma 1. *The equilibrium net earnings of the firm π_t at time t can be written as*

$$\pi_t = \mathcal{A}(K_t) + \mathcal{B}(I_t, K_t) + (F_K K_t - G_K K_t) - (1 + G_I)\, m \left(\int_t^{\infty} e^{-\int_t^s r_\tau d\tau} \left[F_{K_s} K_s - G_{K_s} K_s \right] ds \right) K_t \tag{5}$$

where

$$\mathcal{A}(K) \equiv F(K) - F_K(K)K \tag{6}$$

$$\mathcal{B}(I, K) \equiv G_I(I, K)I + G_K(I, K)K - G(I, K) \tag{7}$$

and m is some strictly increasing function. With convexity degree of F being sufficiently large, $\pi_t < 0$ for small K.

The proof is presented in Appendix C.2. Here, we focus on the underlying intuition. We assume that the adjustment cost function is homogeneous of degree 1, meaning that $\mathcal{B}(I, K) \equiv G_I(I, K)I + G_K(I, K)K - G(I, K) = 0$. This assumption allows us to simplify the firm's net earnings as follows:

$$\pi = \mathcal{A}(K) + \underbrace{(F_K K - G_K K)}_{\text{current benefits}} - (1 + G_I)\, m \left(\int_t^{\infty} e^{-\int_t^s r_\tau d\tau} \underbrace{\left[F_{K_s} K_s - G_{K_s} K_s \right]}_{\text{future benefits}} ds \right) K \tag{8}$$

The first term, $\mathcal{A}(K)$, directly reflects the degree of returns to scale. If F exhibits constant returns to scale, then $\mathcal{A}(K) = 0$. If F is concave, we have $\mathcal{A}(K) > 0$, whereas if F is convex, $\mathcal{A}(K) < 0$. The second term captures the immediate benefit of additional capital, defined as the marginal product of capital minus the marginal adjustment cost associated with additional capital. Similarly, the expression $F_{K_s} K_s - G_{K_s} K_s$ represents the future benefits of additional capital for periods $s > t$. This framework allows us to examine the conditions under which π may turn negative. If F is concave, not only is $\mathcal{A}(K) > 0$, but also the future benefits of additional capital are less than the immediate benefits, meaning firms in this setting lack strong incentives to invest. However, if F is convex, then $\mathcal{A}(K) < 0$, and more importantly, the future benefits of additional capital significantly exceed the immediate benefits. Since m is an increasing function, under certain conditions, the investment demand—driven by discounted future capital benefits—can surpass the immediate benefits. In this case, net earnings become negative. Of course, when K becomes large, the earnings will turn positive again as both G_I and $-G_K$ are decreasing in firm size K.

Based on this interpretation, our hypothesis is that the global increase in firms with negative net earnings is likely to be driven by the growing importance of the convexity component in the corporate production function. In the next section, we present direct evidence supporting this hypothesis.

3.2.3 The dislink between marginal q and average Q

Our model here also implies that when the shape of production function changes, the marginal q deviates from average q. Following Fumio Hayashi (1982), we define a firm's marginal $q \equiv \frac{\partial V}{\partial K}$ and average $Q \equiv \frac{V_t}{K_t}$. Their relationship can be shown as in the following lemma.

Lemma 2. *The relationship between a firm's marginal $q \equiv \frac{\partial V}{\partial K}$ and average $Q \equiv \frac{V_t}{K_t}$ is as follows:*

$$q_t = Q_t - \frac{1}{K_t} \int_t^{\infty} e^{-\int_t^s r_\tau d\tau} \left[\mathcal{A}\left(K_s\right) + \mathcal{B}\left(K_s, I_s\right) \right] ds \tag{9}$$

where $\mathcal{A}$ and $\mathcal{B}$ are the same as in Lemma 1.

The detailed proof is shown in Appendix C.3. Similarly, for simplicity, we assume that the adjustment cost function G is homogeneous of degree one, so that $\mathcal{B}(K, I) = 0$. Under this setup, the relationship between marginal q and average Q is influenced by the shape of the production function F. If F exhibits constant returns to scale, then $\mathcal{A}(K) = 0$, leading to the classical result from Hayashi (1982) that marginal q equals average Q.

However, if F has decreasing returns to scale, then $\mathcal{A}(K) > 0$, resulting in marginal $q <$ average Q. In this situation, the firm's investment demand is low due to the low marginal q. Conversely, if F exhibits increasing returns to scale, then $\mathcal{A}(K) < 0$, resulting in marginal $q >$ average Q. In this case, firms exhibit a strong demand for investment as the high marginal q incentivizes expansion, and average Q becomes less informative for corporate investment when the shape of production function changes over time.

4 Trends in the Shape of Corporate Production Function

4.1 Methodology

In this section, we estimate the shape of the corporate production function and examine its trends over time. Our goal is to impose minimal restrictions on the structure and sequence of inflection points, allowing our methodology to more accurately capture the dynamics present in the data.

Our conjecture is that, for any given firm i, if its total capital remains below a specific threshold or turning point, the firm experiences increasing returns to scale. However, once capital accumulation exceeds this inflection point, the firm shifts to a regime of decreasing returns to scale.[6] The primary econometric challenge is that we cannot observe both returns-to-scale regimes for a single firm simultaneously. Therefore, we treat the firms in each cross-sectional year as random samples drawn along the production function, enabling us to infer these transitions.

Specifically, in each cross-sectional year t from 1980 to 2021, we apply Bayesian MCMC to determine the capital threshold or inflection point, $\overline{K}_t$. This approach allows us to estimate the position of these inflection points for each year, revealing how these thresholds evolve over time. The resulting time series of $\overline{k}_t$ reflects shifts in the inflection points, which provides insight into the evolving nature of corporate production functions over the past several decades.

a two-step estimation strategy The estimation of the inflection point naturally aligns with the broader literature on structural change estimation (e.g., Gregory C Chow, 1960; Richard E Quandt, 1958). A substantial body of research focuses on detecting and testing structural shifts in the relationship between economic variables over time (e.g., Pierre Perron et al., 2006; Alexander Aue and Lajos Horváth, 2013). In cases of abrupt changes, the relationship undergoes a sudden shift once a specific index is surpassed, commonly referred to as the change point.[7] However, in our setting, two challenges arise when applying classical structural change estimation techniques. First, the index variable is continuous rather than a discrete time index. Second, the explanatory variable is subject to endogeneity. To address these issues, we propose a heuristic two-step estimation strategy that integrates Bayesian MCMC to estimate the breakpoint of the continuous capital variable while leveraging the classic Levinsohn and Petrin (2003) method to correct for endogeneity.

In the first step, we tackle the issue posed by the continuous nature of the capital variable. Traditional structural change estimation relies on grid search (for a single break) or dynamic programming as in Jushan Bai and Pierre Perron (2003) (for multiple breaks). These methods become

[6] In our estimation procedure, we do not impose the existence of a threshold or require that the estimated degree of returns to scale in the first component be greater than 1. Instead, we let the data determine these relationships.

[7] A closely related approach is threshold regression (e.g., Howell Tong, 1978; Bruce E Hansen, 2000), which generalizes the threshold variable beyond a time index.

computationally expensive when dealing with large datasets, with a complexity of $O(T^2)$, where T denotes the number of observations. To improve efficiency and precision in estimating the inflection point, we employ the Bayesian MCMC approach proposed by David A Stephens (1994), which is particularly suited for detecting continuous breaks. Following the arguments in Pierre Perron and Yohei Yamamoto (2015), it is straightforward to show that the breakpoint can be consistently estimated even in the presence of potential endogeneity in capital.[8]

In the second step, we divide the data into pre-break and post-break subsamples. While the breakpoint estimation from the first step is unbiased, the slope coefficients remain subject to endogeneity.[9] To address this concern, we construct a rolling 10-year panel by incorporating data from the preceding nine years for each year within each subsample. We then apply the Levinsohn and Petrin (2003) method to correct for endogeneity and iterate forward year by year. As a robustness check, we also implement the Olley and Pakes (1996) approach.[10]

why Bayesian? Two key challenges motivate our use of Bayesian MCMC. First, our model examines the relationship between output and capital, a continuous variable rather than a discrete one. In the literature, structural breaks estimation is often applied to time-series settings with a discrete-time index ($t = 1, 2, \ldots$), where maximum likelihood methods naturally serve as a tool to evaluate model fit when a finite set of possible structural breaks is defined. However, in a continuous context (e.g., capital as in our model), it is impossible to conduct a search over all possible values for the structural break. Bayesian MCMC addresses this issue by allowing us to place continuous distributions, thus facilitating structural break estimation. Technically, Bayesian MCMC also simplifies model fitting, as it is challenging to compute derivatives of the continuous-time (log-)likelihood function, which are required for standard error calculation based on second-order derivatives.

Second, beyond estimating the structural break, we aim to estimate the returns-to-scale parameters. Obtaining the marginal distributions of these parameters is also challenging, as it requires integrating over the distributions of other parameters. Bayesian MCMC enables efficient sampling from the joint posterior distribution by constructing a Markov Chain from the univariate conditional posterior distributions, thus avoiding exhaustive continuous interval searches or

[8] Our simulation results provide further validation of this point. Ping Yu (2015) notes that estimation may be biased when the threshold variable is endogenous. However, as demonstrated by Perron and Yamamoto (2015), the inflection point is affected only in extreme edge cases. As a robustness check, we also implement the Ping Yu and Peter CB Phillips (2018) method and find that our results remain consistent.

[9] We have also conducted simulation studies to further confirm the presence of this endogeneity issue.

[10] Additionally, since we assume that labor choices are determined without frictions, labor does not enter our model directly as the optimal labor choice is an function of existing capital stock. In the absence of labor, the Ackerberg, Caves and Frazer (2015) approach is equivalent to the Levinsohn and Petrin (2003) method.

direct integration over auxiliary parameters. The computational benefits of Bayesian MCMC for such applications are well-documented in the existing literature (e.g., Stephens, 1994).

baseline functional form The functional form used in our baseline estimation is as follows. For a given firm i at time t, let total output be denoted by $Y_{i,t}$ and total capital stock, which includes both tangible and intangible assets, be represented by $K_{i,t}$. We assume a structural change occurs at $\bar{k}_t$, allowing us to model the production function as:

$$Y_{i,t} = \begin{cases} A_t^H \Omega_{i,t} Z_{i,t}^H K_{i,t}^{\alpha_t^H} & \text{if } K_{i,t} < \bar{K}_t \\ A_t^L \Omega_{i,t} Z_{i,t}^L K_{i,t}^{\alpha_t^L} & \text{if } K_{i,t} \geq \bar{K}_t, \end{cases} \tag{10}$$

where A_t^H and A_t^L are common aggregate factors, $\Omega_{i,t}$ denotes unobservable firm-level productivity, $Z_{i,t}^H$ and $Z_{i,t}^L$ are independent log-normal random shocks, and α_t^H and α_t^L represent the returns-to-scale parameters before and after the threshold at time t, respectively. Importantly, we impose *no* restrictions on α^H and α^L (i.e., $\alpha^H > 1$ or $\alpha^L < 1$); instead, the data informs the estimation results of these parameters. Taking the logarithm of both sides and reparameterizing, we derive our baseline estimation model:

$$\log Y_{it} = \begin{cases} a_t^H + \omega_{i,t} + \alpha_t^H \log K_{i,t} + \varepsilon_{i,t}^H & \text{if } K_{i,t} < \bar{K}_t \\ a_t^L + \omega_{i,t} + \alpha_t^L \log K_{i,t} + \varepsilon_{i,t}^L & \text{if } K_{i,t} \geq \bar{K}_t, \end{cases} \tag{11}$$

where $\varepsilon_{i,t}^H \equiv \log Z_{i,t}^H \sim \mathcal{N}(0, \sigma_H^2)$ is independent of $\varepsilon_{i,t}^L \equiv \log Z_{i,t}^L \sim \mathcal{N}(0, \sigma_L^2)$. In addition, $a_{it} \equiv \log A_t$ and $\omega_{it} \equiv \log \Omega_t$. Equation (11) is specified in a cross-sectional framework, with parameter values varying across different years. Within this framework, structural break estimation is conducted by treating capital as a continuous "timing" variable that signals shifts in returns to scale, enabling us to capture structural changes in the production function across different capital levels. As previously mentioned, we utilize the Bayesian MCMC approach to detect the presence of a threshold. To address potential bias in estimating the unobservable $\omega_{i,t}$, we adopt the method proposed by Levinsohn and Petrin (2003) in our baseline analysis. A detailed explanation of the implementation is provided in Section A of the appendix.

4.2 Baseline Evidence

Figure 2 presents long-term changes in the corporate production function using two graphs. Graph (a) shows the evolution of the estimated convexity-concavity threshold over time, while Graph (b) provides time-series plots of the estimated degrees of returns to scale for various production components. This dual perspective allows us to observe both the shifts in capital thresholds and the changing returns-to-scale patterns across different production aspects.

[Figure 2 here]

In Graph (a), the blue curve shows the threshold in natural logarithmic terms, while the red dashed line reflects the raw values in millions of USD. The findings indicate a significant upward trend in this threshold, revealing a steady increase in the capital level $\bar{k}$ at which firms shift from increasing to decreasing returns. Specifically, in 1980, the threshold was estimated at 5.02 thousand dollars; by 2021, it had risen to 1.31 million dollars – a dramatic 261-fold increase. The estimated time trend coefficient, shown in Table 2, is 0.084 and is significant at the 1% confidence level. All these findings suggest that modern firms are experiencing an extended phase of increasing returns to scale, requiring a much larger operational scale before the neoclassical diminishing return effects begin to apply.

In Graph (b), we illustrate the estimated degrees of returns to scale over time. The orange curve, with shaded areas, represents the slope coefficient α^H, while the green curve and its shaded region represent α^L. These estimates are obtained without any prior restrictions on slope magnitudes, providing an unbiased perspective on production dynamics. The shaded areas around each curve show the 95% credible intervals, approximated by two times the posterior standard deviations. The plot reveals two distinct regimes within the aggregate production function. The first component of production function, marked by α^H values consistently above 1 for all the periods after 1986, indicates increasing returns to scale, while the second, represented by α^L, shows values below or close to 1, indicating decreasing returns to scale or a constant one. Despite some fluctuations over time, the average values of α^H and α^L in our sample are approximately 1.11 and 0.99, respectively. Statistical tests further confirm that these two coefficients are significantly different from each other, with a mean of 0.13 and a T-statistics of 21.58, providing strong evidence of these dual phases within the production structure.

The upward trend in the convexity-concavity threshold, along with the distinct patterns of returns to scale in both production components, implies that modern firms operate with a more pronounced phase of initial increasing returns. This trend could be driven by factors such as technological advancements, the scaling of intangible assets, and increased capital intensity. Our model, applied to U.S. public firms, highlights a structural transformation in the production function, aligning with the broader shift towards capital-intensive and intangible-dependent production models (e.g., Maarten De Ridder, 2024; Karabarbounis and Neiman, 2014).

More importantly, this long-term shift in the production function's shape represents a noteworthy development for economic theory, where production functions are central in linking inputs to outputs. Traditional concave functions (e.g., Cobb-Douglas, Constant Elasticity of Substitution) remain widely used in theoretical and empirical work, yet our evidence of evolving production dynamics suggests a need to reconsider standard models to better capture the increasingly capital-heavy and scale-intensive nature of modern production.

[Figure 3 here]

In Graph (a) of Figure 3, we illustrate the percentile position of the estimated threshold over time. To account for the possibility that the upward trend in the threshold may simply reflect increasing firm size, we report the quantile rank of the estimated convexity-concavity threshold for each year. Specifically, for each year t, we compute the relative rank of the convexity-concavity threshold as $\bar{k}_{pc} = \hat{F}_k(\hat{\bar{k}}_t)$, where $\hat{F}_k$ is the empirical cumulative distribution function of total capital. This approach enables us to measure the threshold relative to the overall capital distribution across firms in each year. The results in Graph (a) reveal an upward trend in the quantile rank of the threshold, affirming that the increasing significance of production convexity is not solely due to larger firm sizes. Initially, in 1980, the quantile of the estimated breaking point was approximately 0.11, meaning that only the bottom 11% were in the increasing returns-to-scale phase. By 2021, this quantile has risen to around 0.60, indicating that firms whose capital stocks below 60% of the maximum capital level remain in the increasing returns-to-scale stage. Similarly, the estimated time trend coefficient for this quantile threshold, shown in Table 2, is 0.012 with a T-statistics of 8.707.

To provide further insights into the long-term shifts in the corporate production function, we present Figure 3, which includes two alternative indicators that reveal nuanced aspects of this transformation. In Graph (a), we present an average measure of returns-to-scale, calculated as $\alpha^H \times \bar{k}_{pc} + \alpha^L \times (1 - \bar{k}_{pc})$. This metric represents a weighted average of returns-to-scale across firms, capturing both the increasing returns-to-scale component (denoted by α^H) and the decreasing returns-to-scale component (denoted by α^L), weighted by the relative threshold $\bar{k}_{pc}$. The results in Graph (b) indicate a steady upward trend in the average returns-to-scale over the sample period. At the beginning of the period, the average return-to-scale was 0.97 in 1980. By 2021, this value has risen to 1.07, with a peak of 1.08 in 1999. This suggests that the degree of returns-to-scale has increased by 10% over the past forty years, with the initial phase of increasing returns-to-scale exerting a growing influence on the overall production function of firms in recent decades. Again, the corresponding time trend coefficient is statistically significant at the 1% confidence level.

In addition, Graph (b) provides a measure of the relative importance of convexity versus concavity in production, calculated as the ratio $\frac{\text{convexity}}{\text{concavity}} = \frac{\alpha^H \times \bar{k}_{pc}}{\alpha^L \times (1-\bar{k}_{pc})}$. This ratio offers insights into the dominance of convexity (or increasing returns-to-scale) relative to concavity (or decreasing returns-to-scale) in firms' production processes. The results show a significant rise in the relative importance of convexity throughout the sample period, with an initial value of 0.28 in 1980, climbing to 1.72 by 2021. The estimated time trend coefficient is 0.051 with a T-statistics of 7.729.

Therefore, the evidence presented in Figure 3 complements our findings in Figure 2, highlighting a structural shift in the production function of U.S. public firms. This shift underscores the rising prominence of convexity in production, indicating that firms now operate at a larger scale before reaching diminishing returns.

In addition, the combination of reduced marginal costs and higher fixed costs can naturally lead to a rightward shift in the convex-concave threshold. Intuitively, lower marginal costs make it optimal for firms to produce more before encountering diseconomies of scale, while higher fixed costs incentivize firms to operate at larger scales to spread these costs. A detailed discussion is provided in Appendix B.2.

public v.s. full-sample Due to data limitations, our empirical analysis is based solely on a dataset of public firms. This restriction may affect certain aspects of our findings, particularly the movement of the convexity-concavity threshold and the degree of convexity in the production function. However, it does not impact our primary conclusion regarding the long-run shifts in convexity-concavity thresholds over time. Our analysis suggests that the observed increase in thresholds among public firms likely provides a lower bound for threshold changes across the broader population of firms. This idea is illustrated in Figure A4 in the appendix. Private firms are generally smaller and less likely to have reached the concave portion of their production functions. If public firms are becoming younger—indicating that the IPO threshold has shifted left—this could contribute to the observed trend in the rising degree of returns to scale in the data (as shown in Figure A4 in the appendix). However, as Figure A2 indicates, public firms have been aging over time. From this perspective, the changes in the corporate production function identified in our estimates likely provide a conservative benchmark for the shifts occurring in the broader economy, which includes both public and private firms.

At the same time, this limitation may influence the estimated slope coefficients, particularly for the convex portion of the production function, as seen in Figure A4. Excluding private firms could lead to an underestimation of the degree of convexity, as smaller, non-public firms tend to exhibit stronger increasing returns to scale due to their less mature production processes and lower initial capital requirements. Future research could address this limitation by incorporating private firm data, providing a more comprehensive view of production function dynamics across both public and private firms.

4.3 Robustness Checks

In this section, we conduct a series of additional tests to verify the robustness of our primary conclusions from the baseline analysis. These tests are designed to examine the stability of our findings across various specifications and to assess the sensitivity of our results to alternative assumptions. By implementing these robustness checks, we aim to strengthen the reliability of our core insights regarding the evolving production function dynamics and the long-term shifts in returns to scale.

4.3.1 Simulation

To begin, we assess the efficacy of Bayesian MCMC estimation through simulation studies. This simulation exercise is designed to demonstrate our methodology's ability to accurately detect time-series changes in thresholds. In each cross-section of this example, we set the 30% quantile of total capital as the break point and assign $\alpha^H = 1.3$ and $\alpha^L = 0.7$. Additionally, we allow premature firms to exhibit greater output variability. Specifically, the true model specification in this simulation exercise is as follows:

$$\log y_{i,t} = \begin{cases} -2 + (1.3 + u_t^H) \log k_{i,t} + \varepsilon_{i,t}^H & \text{if } k_{i,t} < \bar{k}_t \\ 2 + (0.7 + u_t^L) \log k_{i,t} + \varepsilon_{i,t}^L & \text{if } k_{i,t} \geq \bar{k}_t, \end{cases} \tag{12}$$

where the independent terms $u_t^H \sim Uniform(-0.25, 0.25)$ and $u_t^L \sim Uniform(-0.1, 0.1)$ introduce time-series variation to the coefficients. The error terms $\varepsilon_{i,t}^H$ and $\varepsilon_{i,t}^L$ are independently distributed with respect to each other, and we assume that $\varepsilon_{i,t}^H \sim \mathcal{N}(\log k_{i,t}, 0.49)$ and $\varepsilon_{i,t}^L \sim \mathcal{N}(\log k_{i,t}, 0.25)$. We model the mean of these error terms to be perfectly driven by the endogenous capital to reflect one possible cause of endogeneity. The coefficients and relative quantiles are set to reflect yearly changes in total capital. Importantly, since we estimate structural breaks independently within each cross-section, this setup more closely mirrors real-world conditions.

Using the simulated dataset, we replicate our previous analysis with the Bayesian MCMC estimation approach. The estimated results for the convexity-concavity threshold and slope coefficients are displayed in Graphs (a) and (b) of Figure A5 in the Appendix. These graphs demonstrate that our methodology accurately identifies the turning point whenever it exists in the data. Notably, the estimated thresholds from our Bayesian MCMC approach closely align with the true values. Additionally, the estimated slope coefficients are nearly identical to the actual values and fluctuate around them over time. Not surprisingly, the slope estimates are biased from the true values of 1.3 and 0.7, respectively. This simulation study, therefore, affirms the reliability and accuracy of the structural break estimation of our Bayesian methodology.

4.3.2 Alternative Functional Form

Afterward, we examine whether our main findings remain consistent when alternative functional forms are used for the convex-concave production function.

continuity First, our baseline model (i.e., Equation (11)) does not assume continuity in the production function, resulting in a potential discontinuity at the turning point $\bar{k}_t$. As a robustness check, we impose continuity at $\bar{k}_t$ by enforcing the condition $a_t^H + \alpha_t^H \log \bar{k}_t = a_t^L + \alpha_t^L \log \bar{k}_t$. Solving for a_t^L and incorporating this restriction into model (11), we can rewrite the equation as follows:

$$\log y_{i,t} = \begin{cases} a_t^H + \omega_{i,t} + \alpha_t^H \log k_{i,t} + \varepsilon_{i,t}^H & \text{if } k_{i,t} < \bar{k}_t \\ a_t^H + \omega_{i,t} + (\alpha_t^H - \alpha_t^L) \log \bar{k}_t + \alpha_t^L \log k_{i,t} + \varepsilon_{i,t}^L & \text{if } k_{i,t} \geq \bar{k}_t. \end{cases} \tag{13}$$

We then replicate our baseline procedures, this time assuming a continuous convex-concave production function. The results for the thresholds and slope coefficients under this continuity constraint are shown in Graphs (a) and (b) of Figure A6. According to these graphs, our main conclusions remain consistent with the modified model specification. We continue to observe an upward trend in the estimated convexity-concavity threshold, which further reinforces the increasing significance of production convexity over time. However, the specific magnitudes show slight deviations from our baseline analysis. The estimated breaking point is approximately 48.1 thousand dollars in 1980, rising to around 2.42 million dollars in 2021 – a nearly 5,000% increase.

Furthermore, the average values of the estimated slope coefficients before and after the threshold across our sample are approximately 1.20 and 0.99, respectively. These estimates differ slightly from the 1.11 and 0.98 obtained with our baseline approach. Despite these minor differences in magnitude, our main conclusions regarding the evolving shape of the production function remain robust. This robustness indicates that the observed shifts in production convexity reflect a fundamental change, largely unaffected by variations in the specific model assumptions applied.

exponential functional form Another commonly used functional form in the convex-concave production function literature is the S-shaped exponential form (e.g., Skiba, 1978). Specifically, the functional form applied in our robustness check is as follows:

$$y_{i,t} = \frac{A_t \Omega_{i,t} Z_{i,t}}{1 + e^{-\gamma_t (\log k_{i,t} - \bar{k}_t)}}, \tag{14}$$

where $\log Z_{it}$ follows an iid normal distribution $\mathcal{N}(0, \sigma^2)$. The definitions of y and k remain the same as in our baseline model (Equation (11)). Taking the logarithm of both sides and redefining

$a_{it} \equiv \log A_t$, $\omega_{it} \equiv \log \Omega_t$, and $\varepsilon_{it} \equiv \log Z_{it}$, we get the following equation:

$$\log y_{i,t} = a_t + \omega_{i,t} - \log\left(1 + \exp\left\{\gamma_t \bar{k}_t - \gamma_t \log k_{i,t}\right\}\right) + \varepsilon_{i,t}. \tag{15}$$

Compared to the functional form in our baseline model, this S-shaped exponential form is smooth and continuously differentiable at every point, making it particularly advantageous for theoretical studies. For this model, we estimate the parameters using the maximum likelihood method due to its straightforward form. The standard errors and t-values of the estimators are derived from the inverted observed Fisher information matrix. The estimated results for the convexity-concavity threshold ($\bar{k}_t$) and slope coefficient (γ_t) are shown in Graphs (c) and (d) of Figure A6, respectively.

These graphs demonstrate that the upward trend in the estimated convexity-concavity threshold persists even with the S-shaped functional form. The average estimated breaking point is approximately 67.15 million dollars in 1980s, increasing to around 254.30 million dollars in 2021, marking a 379% increase, which aligns with our findings in the baseline analysis. Furthermore, the average value of the coefficient γ is estimated to be around 1.11 across the sample. These results reinforce our primary conclusions regarding the growing importance of production convexity over time, suggesting robustness to alternative functional forms.

4.3.3 Alternative TFP Estimation Method

In this robustness check, we apply the Olley and Pakes (1996) method for TFP estimation in place of the Levinsohn and Petrin (2003) endogeneity correction used in our baseline analysis. The Olley-Pakes approach addresses endogeneity by employing investment as a proxy variable to control for unobserved productivity shocks, offering an alternative means of mitigating potential biases in production function estimates. For this exercise, we use total investment in both tangible and intangible capital, as well as firm age.

Figure A7 in the Appendix presents the results using this adjustment. In Graph (a), we show the estimated degrees of returns-to-scale (again, α^H for the increasing returns phase and α^L for the decreasing returns phase) over time. Consistent with the baseline results, we find that α^H generally remains above 1, while α^L stays close to 1, reflecting the two-phase structure of returns to scale among firms. Throughout our sample, the average values for these two parameters are 1.28 and 1.03, respectively. The shaded areas denote confidence intervals, indicating that both are statistically above 1. The T-statistics for α^H and α^L are 13.20 and 3.82, respectively. Compared with our estimation results using Levinsohn and Petrin (2003) method, the estimated α^H values are relatively large.

Similar to our baseline findings, Graph (b) displays the average measure of returns to scale. This figure shows a clear upward trend, especially after the early 2000s, underscoring the growing importance of the convex portion of the production function. The average return-to-scale was 0.98 in 1980, and it has increased to 1.20 in 2021. Graph (c) illustrates the ratio of convexity to concavity. Our calculated ratio steadily rises over the sample period, rising from 0.10 in 1980 to 2.03 in 2021, indicating a shift toward convexity in firms' production functions. This pattern aligns with our main findings, affirming the robustness of the results under an alternative TFP estimation method.

4.3.4 Alternative Econometric Approach

Yu and Phillips (2018) introduce a nonparametric method for estimating breakpoints that does not rely on instrumental variables. Their approach leverages information about structural changes to identify abrupt shifts. However, this method involves nonparametric screening across all possible breakpoints, leading to a computational cost that significantly exceeds that of our approach. The computational burden is particularly prohibitive for detecting breaks in a reasonable time when the sample size is large.[11]

As a robustness check, we implement their method by discretizing total capital into 200 grids to approximate the true breakpoints. The corresponding results are shown in Figure A8 in the appendix. The findings confirm that our key conclusions remain robust under this alternative econometric approach. Specifically, there is an upward trend in the estimated threshold for the relationship between output and input, with the estimated degrees of returns to scale exceeding 1 before the threshold and falling below 1 afterward. However, the estimated magnitudes differ. The breaking point is estimated at approximately \$76,500 in 1980, rising to about \$204,000 in 2021—a nearly 267% increase. The average returns to scale are 1.12 before the threshold and 0.98 after the threshold. Therefore, our key conclusions remain robust when using this alternative econometric approach.

4.4 Industry-Level Evidence

In this section, we examine industry-level evidence of long-run changes in corporate production functions. Specifically, we analyze the estimated convexity-concavity threshold in level, threshold in percentile, average return-to-scale, and the relative importance of convexity to concavity across various industries. These measures are estimated and constructed in the same way as in our baseline analysis, enabling a comparison of industry-specific trends against aggregate findings. We

[11] Our rough estimate suggests their method is approximately 144 times slower than ours.

present the average return-to-scale results for 10 different Fama-French industries in Figure 4, with results for the other three measures shown in Figure A9 in the appendix. Additionally, the estimated time-trend coefficients, displayed in Table 2, summarize the sign and significance of time trends for each measure within each industry.

[Figure 4 here]

The first key observation from our analysis is that the patterns of change differ substantially across industries. For instance, industries such as *Manufacturing* and *Healthcare* exhibit significant increases in both the baseline threshold and average return-to-scale over time, suggesting a shift towards greater economies of scale and an increasing role of convexity. In the 1980s, the average return-to-scale in *Manufacturing* was only 0.92, which rose to 1.05 by the 2010s. Similarly, *Healthcare* saw an increase from 0.72 to 1.05 over the past forty years. A modest increase is also observed in *Business Equipment*, where the average return-to-scale grew from 1.02 to 1.05, peaking at 1.12 around 2000. Certain industries, such as *Consumer Nondurables, Consumer Durables, Telephone and Television*, and *Other*, display a U-shaped pattern, with initial declines followed by increases in return-to-scale. In contrast, *Wholesale and Retail* exhibits a downward trend in the convexity-concavity threshold, decreasing from 1.07 to 0.98 over the past few decades, indicating a decline in economies of scale for this sector. For other sectors, such as *Energy* and *Utilities*, no clear trends emerge, likely due to more pronounced business cycle effects.

[Table 2 here]

To further categorize industries, we classify an industry as having increasing returns-to-scale if its time-trend estimate for return-to-scale is positive and statistically significant. Conversely, an industry is classified as having a negative trend if the estimate is negative and significant, or as having no trend if the estimate is insignificant. Based on the time-trend estimates in Table 2, industries with the strongest upward trends include *Consumer Nondurables* (0.002), *Manufacturing* (0.005), *Healthcare* (0.006), and *Telephone and Television* (0.006), all exceeding the coefficient estimated for the aggregate economy (0.001). These industries show significant positive time-trend coefficients across multiple measures, highlighting an increasing emphasis on convexity in their production functions. In contrast, industries such as *Consumer Durables* (-0.002), *Wholesale and Retail* (-0.004), and *Other* (-0.0008) display negative or insignificant trends, suggesting either a decline or stability in returns-to-scale characteristics. These results align with findings from Kariel, Savagar and Mainente (2022), who, using UK data, report significant variability in returns-to-scale across sectors. For instance, manufacturing consistently exhibits high returns-to-scale, with a positive trend over time, whereas sectors such as wholesale, trade, and transport display returns close to constant returns-to-scale.

As illustrated in Figure A9 and reflected in the time-trend coefficients in Table 2, using alternative indicators reveals slight differences in industry-specific trends. For example, in the *Consumer Durables* sector, we observe a significant upward trend in both the threshold level and percentile, though not in the average return-to-scale. These differences may stem from measurement noise or varying business cycle characteristics across industries. Additionally, the U-shaped pattern seen in industries such as *Telephone and Television* suggests that focusing on only the most recent decade may yield different conclusions. However, for industries like *Consumer Nondurables, Manufacturing*, and *Healthcare*, the evidence of structural shifts towards higher returns-to-scale and dominance of convexity remains clear and consistent, with all four indicators showing a significant upward trend.

In summary, our industry-level analysis reveals that shifts in production functions are not uniform across sectors. Differences in time-trend estimates underscore the unique trajectories of each industry, highlighting the evolving landscape of production technologies and economies of scale shaped by industry-specific dynamics. This sectoral diversity reflects how structural factors, technological progress, and market conditions interact differently within each industry, leading to varying paths of production function development.

4.5 Global Evidence

Beyond the industry-level analysis, we extend our investigation to country-level evidence of long-run changes in corporate production functions to evaluate whether our main findings hold across diverse economic settings. Following a similar approach to the industry-level analysis, we examine the baseline threshold, threshold in percentile, average return-to-scale, and the relative importance of convexity to concavity for each country. By maintaining consistency with these measures from our baseline analysis, we ensure that country-specific variations in production function characteristics over time are comparable. Table 2 provides a summary of the time-trend estimates for each country, highlighting the direction and statistical significance of changes in these core metrics. The average return-to-scale results for these countries are presented in Figure 5.

[Figure 5 here]

Our findings reveal heterogeneous trends across countries, reflecting differing economic structures, industrial compositions, and stages of development. For instance, countries such as *China* and *India* show substantial increases in both the baseline threshold and average return-to-scale, suggesting a trend toward greater economies of scale and a more prominent role of convexity in production. Specifically, the average return-to-scale in China has increased from 0.75 in 2000 to 1.03 in 2019, while India shows a rise from 0.64 to 0.94 over the same period. These changes align with rapid economic growth and increased capital intensity in these nations, which may reflect shifts toward larger-scale and more capital-intensive production processes. In contrast, countries such as *Canada* and *Japan* display relatively stable or even slightly declining trends in these measures, suggesting a slower or more conservative evolution of production structures within these economies over the sample period. *South Korea* and the *UK* exhibit a U-shaped trend, with an initial decline in return-to-scale followed by a moderate upward adjustment in recent years. This pattern may indicate structural adaptation in response to evolving economic demands.

To further categorize countries, we define a country as experiencing increasing returns-to-scale if its time-trend estimate for return-to-scale is positive and statistically significant. Conversely, a country may exhibit a negative trend if the estimate is negative and significant, or show no trend if the estimate is insignificant. Based on the time-trend estimates, countries with the most pronounced upward trends include *China* (0.016), *India* (0.017), *Germany* (0.003), *Thailand* (0.006), *Japan* (0.002), and *Australia* (0.027). In contrast, countries such as *South Korea* (-0.004) and *Sweden* (-0.011) exhibit negative trends, indicating a long-term decline in returns-to-scale characteristics. For other countries, including *Canada*, *France*, and the *UK*, the estimated time trends are statistically insignificant, suggesting that changes in production structure are more subdued in these economies.

As seen in Table 2, alternative indicators reveal slight differences in country-specific trends, yet the overarching conclusion underscores the diversity in production function dynamics across various economic contexts. Our analysis here highlights the influence of country-specific factors—such as economic policies, resource allocation, industrial composition, and stages of development—in shaping the evolution of corporate production functions. Countries like *China* and *India* appear to be undergoing substantial structural transformations, shifting toward higher returns-to-scale and an increased emphasis on convexity, reflecting a move toward increasingly capital-intensive and scale-driven production models. In contrast, other nations maintain more stable production structures, possibly due to factors such as regulatory environments, market saturation, or economic policy stability. These findings underscore how production dynamics are shaped by a complex interplay of economic and institutional factors that vary widely across global contexts, suggesting that shifts in production function characteristics are profoundly influenced by local economic conditions and policy frameworks.

4.6 Reduced-form Evidence

Before proceeding to our quantitative exercise, we investigate whether industries with higher convexity-concavity thresholds are associated with a higher fraction of unprofitable firms. To ensure precise estimation of the production function, we use the Fama-French 10 industry classifications, which provide sufficiently large sample sizes for each industry.

Figure 6 displays a Binscatter plot illustrating the relationship between the industry-level convexity-concavity threshold and the proportion of firms with negative net earnings. The red dashed line represents the linear-fit regression line. The figure reveals a clear and positive relationship: industries with higher convexity-concavity thresholds are more likely to have a larger share of firms reporting negative net earnings. This positive association, observed in the reduced-form regression, supports our hypothesis that a higher threshold aligns with an increased prevalence of unprofitability within an industry.

[Figure 6 here]

Table 1 further substantiates this finding through regression results. Across various model specifications, the positive relationship between the convexity-concavity threshold and the share of unprofitable firms remains robust, even when controlling for year and industry fixed effects. In terms of economic significance, our results indicate that a one standard deviation increase in the convexity-concavity threshold (2.47 units) corresponds to a 1.73 to 4.69 percentage point increase in the proportion of unprofitable firms. This finding translates to a 0.10 to 0.27 standard deviation increase in the share of firms with negative earnings.

[Table 1 here]

These findings imply that the shape of the production function has economic relevance in determining the profitability landscape within an industry. Industries with higher convexity-concavity thresholds—likely driven by advancements in transformative technologies such as digitization—experience shifting economies of scale that affect business models, intensify competition, and increase the prevalence of unprofitable firms within these sectors. As industries adopt more capital-intensive and scalable technologies, the initial phase of production expansion extends, leading to higher upfront costs and potentially delayed profitability. This pattern underscores the broader influence of production function shape on industry dynamics and profitability distributions, suggesting that changes in production structure can impact firms' financial outcomes and competitive environments.

5 Quantitative Analysis

5.1 Model Framework

Our model framework builds on the firm dynamics model of Hopenhayn (1992), with modifications to include capital accumulation and adjustment costs. A detailed setup of the model can be found in Section D in the appendix. In this framework, firms own their capital stock and make investment decisions under both convex and non-convex adjustment costs. The equilibrium focuses on optimal decision-making regarding entry, exit, and investment strategies.

A key difference in our approach is the use of a convex-concave production function rather than the commonly applied Cobb-Douglas form. The production function is specified as follows:

$$y_{i,t} = z_{i,t} \left(k_{i,t}^{\beta} l_{i,t}^{1-\beta} \right)^{\alpha_t^H \times \mathbb{I}_{k_{i,t} < \bar{k}_t} + \alpha_t^L \times \mathbb{I}_{k_{i,t} \geq \bar{k}_t}} \tag{16}$$

where $y_{i,t}$ represents output, $k_{i,t}$ is capital, $l_{i,t}$ is labor, and $z_{i,t}$ is individual firm productivity. The parameters α_t^H and α_t^L represent the degree of returns-to-scale before and after the threshold $\bar{k}_t$, respectively. The indicator function $\mathbb{I}_{k_{i,t} < \bar{k}_t}$ or $\mathbb{I}_{k_{i,t} \geq \bar{k}_t}$ activates the relevant returns-to-scale parameter based on the capital level relative to the threshold. β represents the usual capital share, and labor can be frictionless chosen.

Importantly, these returns-to-scale parameters (α_t^H, α_t^L) and the threshold $\bar{k}_t$ are time-varying and are calibrated based on our empirical estimates in Section 3 using the baseline model specification. This approach enables our model to capture dynamic shifts in production function characteristics, allowing for a nuanced analysis of firms' capital accumulation decisions under changing economic conditions.

5.2 Quantitative Performance

Graph (a) of Figure 7 presents a comparison between the model's predictions and actual data on the share of firms with negative earnings over time. As shown, our model successfully captures both the general upward trend and fluctuations observed in the data, indicating its effectiveness in replicating key dynamics in the proportion of unprofitable firms. The correlation between the model and data counterparts is 0.74, statistically significant at the 1% level. The share of unprofitable firms exhibits cyclical patterns, with notable increases during economic downturns, such as in the early 2000s and the financial crisis of 2008. This cyclical behavior aligns well with economic conditions, as downturns typically result in reduced revenues and increased financial strain, pushing some firms into negative earnings. In our model, this outcome emerges from a substantial decline in the estimated degree of returns-to-scale, α^H, around the early 2000s.

To gain further insights into the underlying mechanisms, Figure 8 presents the equilibrium value function, investment decisions, and earnings for firms across varying productivity levels. The subscript of z in the figure indicates the relative level of productivity. As seen in Figure 8, Graph (b), firms with low capital stocks but high productivity tend to invest heavily in capital accumulation. This behavior reflects their intention to rapidly scale production and secure future returns. However, due to the convex nature of adjustment costs, the marginal costs of investment rise with increasing capital levels. Consequently, firms with lower initial capital but strong growth potential are more prone to experience negative earnings in the initial stages, primarily due to the burden of adjustment costs. Graph (c) illustrates this effect, showing that this initial investment surge imposes high costs that can exceed revenues, resulting in negative earnings, particularly for young and productive firms that are building their capital base. Intuitively, these firms tolerate short-term losses as part of their strategy to achieve future profitability, anticipating that as they accumulate capital, their need for substantial investment will decrease, allowing them to shift toward profitability as adjustment costs diminish and production efficiency improves.

In summary, our model demonstrates a strong quantitative capability in replicating the observed share of unprofitable firms across various economic cycles. This suggests that the evolving shape of the corporate production function likely accounts for the rising prevalence of unprofitable firms over the past few decades.

5.3 Implications on Corporate Market Power

Our analysis contributes to the ongoing debate over the evolution of corporate market power in the U.S. Recent studies have used different methodologies to estimate markups, which can be broadly categorized into two main approaches. The first is the *demand-system approach*, which estimates product-level markups by analyzing price and quantity data under the assumption of Nash-Bertrand competition. This method is widely applied to consumer goods, using datasets such as NielsenIQ Retail Scanner Data, as seen in the works of Hendrik Döpper, Alexander MacKay, Nathan Miller and Joel Stiebale (2024); James Brand (2021); Enghin Atalay, Erika Frost, Alan T. Sorensen, Christopher J. Sullivan and Wanjia Zhu (2023). Since scanner data provide detailed information on unit sales and revenues at the universal product code (UPC), store, and weekly levels for retail chains, this approach offers granular insights into pricing behavior. However, a key limitation is its narrow focus on tangible consumer goods, which may not capture broader trends in market power. This concern is amplified by a growing body of research emphasizing the rising importance of the intangible economy—such as intellectual property, digital platforms, and brand value—over the past few decades (e.g., Carol Corrado, Jonathan Haskel, Cecilia Jona-

Lasinio and Massimiliano Iommi, 2022; Nicolas Crouzet, Janice C. Eberly, Andrea L. Eisfeldt and Dimitris Papanikolaou, 2022; De Ridder, 2024; Andrea Chiavari and Sampreet Goraya, 2021). As a result, questions have been raised about the representativeness of NielsenIQ data for assessing economy-wide shifts in market power.

The second approach is the *production function method*, which aims to infer markups from revenue data by estimating the quantity-based production function under the assumption that firms equate marginal revenue product with marginal input cost. While this method provides a useful lens, it also has several limitations. First, as highlighted in the misallocation literature (e.g., Diego Restuccia and Richard Rogerson, 2017; Chang-Tai Hsieh and Peter J. Klenow, 2009; Joel M. David and Venky Venkateswaran, 2019), observed changes in the marginal product of inputs may not necessarily reflect true shifts in productivity. Instead, they could indicate time-varying distortions in input allocation. Second, this method does not fully disentangle the effects of market power from those driven by technological change. For example, De Loecker, Eeckhout and Unger (2020) rely on the following formulation for estimating markups:

$$markup^{QJE}it = \theta^v_{i,t} \frac{P_{i,t}Q_{i,t}}{P^v_{i,t}v_{i,t}}$$

where $\theta^v_{i,t}$ denotes the output elasticity of the flexible input v. According to this expression, markups can increase either due to higher output prices ($P_{i,t}$) or due to declining marginal input costs ($P^v_{i,t}$). Thus, the observed markup trends could be driven by rising corporate market power, or alternatively, by enhanced returns to scale—microfounded in our model as reduced variable costs and increased fixed costs, as elaborated in Appendix B.1.

As discussed in Section 3.2.1, our estimated α_t captures a combination of influences: changes in corporate market power (ε_t), adjustments in returns to scale for quantity production (s_t), and shifts in factor shares (γ_t). Specifically, an increase in returns to scale (higher s_t) makes the revenue production function more convex, raising α_t, while greater market power (lower ε_t) makes it more concave, reducing α_t. This interplay implies that even if corporate market power is rising, its impact on α_t may be moderated or even masked by simultaneous technological advances.

Our model simulations, depicted in Panel (b) of Figure 7, provide key insights. When attributing the entire observed rise in α_t solely to technological change—without any increase in pricing power—the estimated average markup (blue line) still rises over time, accounting for a significant share of the empirical upward trend.[12] More specifically, in our quantitative analysis, techno-

[12] Importantly, the upward trajectory of the estimated markup should not be equated with growing market power. Our model also reveals a notable negative relationship between estimated markups and firms' incidence of negative earnings. Highly productive, low-capital firms tend to invest heavily, often resulting in negative earnings. As shown by the orange line in Figure A10 (appendix), the cross-sectional correlation between net earnings-to-sales ratios and esti-

logical factors account for approximately 42.9% of the markup increase between 1980 and 2016. However, notable discrepancies remain between our model's predictions and empirical trends. For instance, while our model attributes most of the markup growth to the 1990s, empirical data indicate that the trend began earlier, around the 1980s. Moreover, while our technology-driven simulations cap markups at around 1.40, observed markups reach as high as 1.60. These discrepancies point to additional drivers beyond technology—such as changes in competitive dynamics, regulatory shifts, or firm-level strategic behavior.

This becomes even clearer when examining markup percentiles, as shown in Panel (c) of Figure 7. Under the assumption that all changes stem from technology, our estimates show rising markups across percentiles, with the sharpest increase at the 90th percentile. Yet, the patterns are similar across the distribution, reflecting uniform exposure to technological change. This contrasts with findings by De Loecker, Eeckhout and Unger (2020). As seen in Graph (d), in the data, lower-percentile markups remain stable over time, meanwhile, the 90th percentile experiences a dramatic rise to levels as high as 2.6. This divergence strongly suggests that factors beyond technology—including rising market concentration, shifts in strategic pricing, and regulatory changes—are integral to understanding the full scope of markup dynamics.

Our analysis complements the institutional explanation advanced by Thomas Philippon (2019), who attributes rising markups and market concentration in the U.S. to regulatory choices, contrasting them with European patterns. Philippon argues that in industries such as telecoms, airlines, and broadband, European consumers enjoy better prices and services despite similar technologies, thanks to stricter regulatory oversight. This suggests that policy differences, rather than technological determinism, largely drive the divergence. However, if economies of scale and technological innovation explain a substantial share of the markup rise, Europe's cautious regulatory approach—while successful in preserving competition—may inadvertently inhibit transformative innovations. Our findings thus shift the focus from a pure concentration story to one where technological forces could be a dominant factor, calling for policy strategies that balance efficiency gains from technological change with the need to curb excessive market power.

6 Conclusion

Using firm-level data on the accounts of all publicly traded firms, our study documents a significant shift in the shape of corporate production functions since the 1980s, moving from the tra-

mated markups has shifted from positive to negative over the past five decades. Post-1980s, this correlation stabilizes at around -0.1 and is statistically significant at the 99% confidence level. This trend is closely tied to the rising importance of intangible capital investment, proxied by net XGSA.

ditional concave form to a sigmoidal (convex-concave) structure. This evolution, characterized by an increasing importance of the convexity component, challenges the conventional assumptions of firm production behavior embedded in standard macroeconomic models. Our analysis, conducted across a wide range of industries and countries, suggests that this transformation is not limited to specific sectors or regions but represents a broader, global change in production dynamics.

Our findings open several avenues for future research. First, further studies could extend this analysis to private firms, as our current dataset is limited to publicly traded companies. Private firms, which tend to be smaller and less capital-intensive, may exhibit different production dynamics, and including them could provide a more comprehensive picture of the evolution in production functions. Second, while our study focuses on the convexity-concavity structure of production, future research could investigate how these changes impact other firm behaviors, such as innovation, investment in intangibles, and competitive strategies. Understanding how convex production environments influence firms' strategic choices could yield insights into broader economic trends, including market concentration and the rise of superstar firms. Third, future research could explore the role of economic policies and institutional factors in shaping the observed shifts in production function shapes. Investigating how tax policies, labor regulations, and access to capital influence the adoption of convex production technologies across different countries could provide policymakers with valuable tools to manage economic growth and promote efficient resource allocation. Lastly, while our model incorporates capital adjustment costs, future studies could develop more nuanced frameworks that account for other forms of adjustment frictions, such as labor mobility or entry-exit barriers. Incorporating these complexities could yield a deeper understanding of the welfare implications associated with the changing production environment, particularly in how these factors affect the cyclical dynamics of firm profitability and market structure.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." *Econometrica,* 83(6): 2411–2451.

Akao, Ken-Ichi, Takashi Kamihigashi, and Kazuo Nishimura. 2011. "Monotonicity and Continuity of the Critical Capital Stock in the Dechert–Nishimura Model." *Journal of Mathematical Economics,* 47: 677–682.

Askenazy, Philippe, and Cuong Le Van. 1999. "A Model of Optimal Growth Strategy." *Journal of Economic Theory,* 85: 24–51.

Atalay, Enghin, Erika Frost, Alan T. Sorensen, Christopher J. Sullivan, and Wanjia Zhu. 2023. "Scalable Demand and Markups." National Bureau of Economic Research Working Paper w31230.

Aue, Alexander, and Lajos Horváth. 2013. "Structural breaks in time series." *Journal of Time Series Analysis,* 34(1): 1–16.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *Quarterly Journal of Economics,* 135(2): 645–709.

Azariadis, Costas, and Allan Drazen. 1990. "Threshold Externalities in Economic Development." *Quarterly Journal of Economics,* 105: 501–526.

Bai, Jushan, and Pierre Perron. 2003. "Computation and analysis of multiple structural change models." *Journal of applied econometrics,* 18(1): 1–22.

Bond, Steve, and Giulio Gottardo. 2024. "Semi-Parametric Bounds for Weighted Harmonic Means of Price-Cost Markups." Working Paper.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. 2021. "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data." *Journal of Monetary Economics,* 121: 1–14.

Brand, James. 2021. "Differences in Differentiation: Rising Variety and Markups in Retail Food Stores." Available at SSRN 3712513.

Carlin, Bradley P, Alan E Gelfand, and Adrian FM Smith. 1992. "Hierarchical Bayesian analysis of changepoint problems." *Journal of the royal statistical society: series C (applied statistics),* 41(2): 389–405.

Chiavari, Andrea, and Sampreet Goraya. 2021. "The Rise of Intangible Capital and the Macroeconomic Implications."

Chow, Gregory C. 1960. "Tests of equality between sets of coefficients in two linear regressions." *Econometrica: Journal of the Econometric Society,* 591–605.

Cobb, Charles W., and Paul H. Douglas. 1928. "A Theory of Production." *American Economic Review,* 18(Supplement): 139–165.

Corrado, Carol, Jonathan Haskel, Cecilia Jona-Lasinio, and Massimiliano Iommi. 2022. "Intangible Capital and Modern Economies." *Journal of Economic Perspectives,* 36(3): 3–28.

Crouzet, Nicolas, Janice C. Eberly, Andrea L. Eisfeldt, and Dimitris Papanikolaou. 2022. "The Economics of Intangible Capital." *Journal of Economic Perspectives,* 36(3): 29–52.

David, Joel M., and Venky Venkateswaran. 2019. "The Sources of Capital Misallocation." *American Economic Review,* 109(7): 2531–2567.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *Quarterly Journal of Economics,* 135(2): 561–644.

De Ridder, Maarten. 2024. "Market Power and Innovation in the Intangible Economy." *American Economic Review,* 114(1): 199–251.

Döpper, Hendrik, Alexander MacKay, Nathan Miller, and Joel Stiebale. 2024. "Rising Markups and the Role of Consumer Preferences." National Bureau of Economic Research Working Paper w32739.

Eisfeldt, Andrea L., and Dimitris Papanikolaou. 2013. "Organization capital and the cross-section of expected returns." *The Journal of Finance,* 68(4): 1365–1406.

Gao, Chengyu, and Matthias Kehrig. 2021. "Returns to scale and productivity dispersion in US manufacturing." *American Economic Journal: Macroeconomics,* 13(1): 101–137.

Ginsberg, William. 1974. "The multiplant firm with increasing returns to scale." *Journal of Economic Theory,* 9: 283–292.

Griliches, Zvi. 1957. "Hybrid corn: An exploration in the economics of technological change." *Econometrica,* 25(4): 501–522.

Gutierrez, German, and Thomas Philippon. 2017. "Investmentless Growth: An Empirical Investigation." *Brookings Papers on Economic Activity,* 48: 89–190.

Hall, Robert E., and Charles I. Jones. 1999. "Why do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics,* 114(1): 83–116.

Hansen, Bruce E. 2000. "Sample splitting and threshold estimation." *Econometrica,* 68(3): 575–603.

Hayashi, Fumio. 1982. "Tobin's Marginal q and Average q: A Neoclassical Interpretation." *Econometrica,* 50(1): 213–224.

Hoberg, Gerard, and Gordon Phillips. 2021. "Scope, Scale and Competition: The 21st Century Firm." Tuck School of Business Working Paper No. 3746660.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5): 1127–1150.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Jones, Callum, and Thomas Philippon. 2016. "The Secular Stagnation of Investment."

Karabarbounis, Loukas, and Brent Neiman. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics*, 129(1): 61–103.

Kariel, Joel, Anthony Savagar, and Joao Mainente. 2022. "Returns to Scale in the UK." *Working Papers*.

Lashkari, Danial, Max Bauer, and Vianney Boussard. 2019. "Returns to scale and the impact of information technology on productivity: Evidence from French firms." *Quarterly Journal of Economics*, 134(3): 1099–1152.

Levinsohn, James, and Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *The Review of Economic Studies*, 70(2): 317–341.

Li, Erica X.N., Guoliang Ma, Shujing Wang, and Cindy Yu. 2021. "Fundamental anomalies." *Available at SSRN 3783526*.

Li, Erica X.N., Haitao Li, Shujing Wang, and Cindy Yu. 2019. "Macroeconomic risks and asset pricing: Evidence from a dynamic stochastic general equilibrium model." *Management Science*, 65(8): 3585–3604.

Olley, G Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–1297.

Perron, Pierre, and Yohei Yamamoto. 2015. "Using OLS to estimate and test for structural changes in models with endogenous regressors." *Journal of Applied Econometrics*, 30(1): 119–144.

Perron, Pierre, et al. 2006. "Dealing with structural breaks." *Palgrave handbook of econometrics*, 1(2): 278–352.

Peters, Ryan H., and Lucian A. Taylor. 2017. "Intangible Capital and the Investment-q Relation." *Journal of Monetary Economics*, 123(2): 251–272.

Philippon, Thomas. 2019. *The Great Reversal: How America Gave Up on Free Markets.* Cambridge, MA:Harvard University Press.

Quandt, Richard E. 1958. "The estimation of the parameters of a linear regression system obeying two separate regimes." *Journal of the american statistical association*, 53(284): 873–880.

Restuccia, Diego, and Richard Rogerson. 2017. "The Causes and Costs of Misallocation." *Journal of Economic Perspectives*, 31(3): 151–174.

Ruzic, Dimitrije, and Kwan Yeung Ho. 2019. "The evolution of returns to scale in U.S. manufacturing: A reexamination." *Journal of Economic Dynamics and Control,* 104: 105783.

Samuelson, Paul. 1979. "Paul Douglas's Measurement of Production Functions and Marginal Productivities." *Journal of Political Economy,* 87(5): 923–939.

Skiba, A. K. 1978. "Optimal Growth with a Convex-Concave Production Function." *Econometrica,* 46(3): 527–539.

Solow, Robert. 1957. "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics,* 39(3): 312–320.

Stephens, David A. 1994. "Bayesian retrospective multiple-changepoint identification." *Journal of the Royal Statistical Society: Series C (Applied Statistics),* 43(1): 159–178.

Tong, Howell. 1978. "On a threshold model." *Pattern recognition and signal processing,* 575–586.

Traina, James. 2021. "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements." Unpublished Working Paper.

Wooldridge, Jeffrey M. 2009. "On estimating firm-level production functions using proxy variables to control for unobservables." *Economics Letters,* 104(3): 112–114.

Yu, Ping. 2015. "Consistency of the least squares estimator in threshold regression with endogeneity." *Economics Letters,* 131: 41–46.

Yu, Ping, and Peter CB Phillips. 2018. "Threshold regression with endogeneity." *Journal of Econometrics,* 203(1): 50–68.

Figure 1: **The Rise of Firms with Negative Net Earnings**

Notes: Graph (A) presents the time-series plot of the fraction of unprofitable public firms. In each year, we count the number of firms with negative profits and divide it by the total number of firms. We use two different profitability measures – gross profits (Compustat data item *GP*) and net earnings (Compustat data item *NI*) – and two different aggregating approaches – weighted and unweighted. The weight is computed as the economy's output share of the industry that a firm belongs to. Data is obtained from *Compustat*. Graph (B) presents the time-series plot of the fraction of unprofitable IPOs. In each year, we count the number of IPOs with negative net earnings and divide it by the total number of IPOs. The information related to corporate earnings is measured at the most recent twelve months before going public. The share of IT stocks is computed as the relative ratio of IT-related IPOs to total IPOs in each year. Data is obtained from Jay Ritter's personal website. Graph (C) presents the time-series plot of the fraction of unprofitable public firms for a global dataset *Compustat Global*. We adopt the same measures as Panel (A). Graph (D) presents the binscatter plot between the fraction of firms with negative net earnings and log real GDP per capita across different countries. Real GDP per capita is obtained from the Penn World Table (PWT) and computed as output-side constant-price real GDP divided by total population.

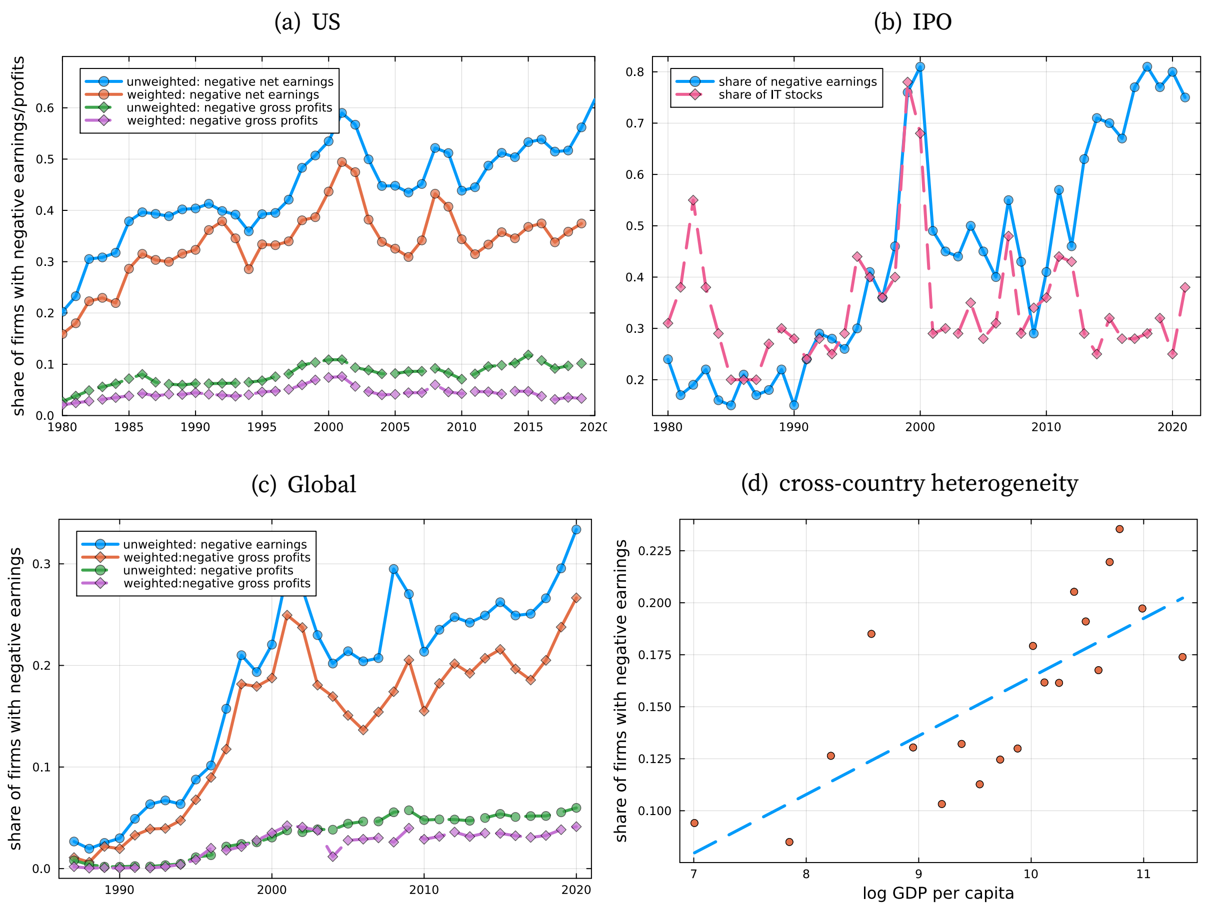


Figure 2: **Long-run Changes in Corporate Production Function: Baseline Evidence**

Notes: Graph (a) shows the estimated cross-sectional convexity-concavity threshold using Bayesian Markov Chain Monte Carlo (MCMC) estimation method. The blue solid line indicates the time-series in natural log scale, while the red dashed line shows the same values in million U.S. dollars. Graph (b) shows the estimated degree of return-to-scale, and the orange and green lines indicate α_t^H and α_t^L, respectively. The band shows the 95% credible interval approximated with two times posterior standard deviations. The baseline model specification is shown in Equation (11) with y being total output (Compustat data item *SALE*) and k the sum of physical capital (Compustat data item *PPENT*) and intangible capital measured by Peters and Taylor (2017). We correct for the possible endogeneity issue by using Levinsohn and Petrin (2003) approach. Data is obtained from *Compustat*.

(a) threshold

(b) return-to-scale

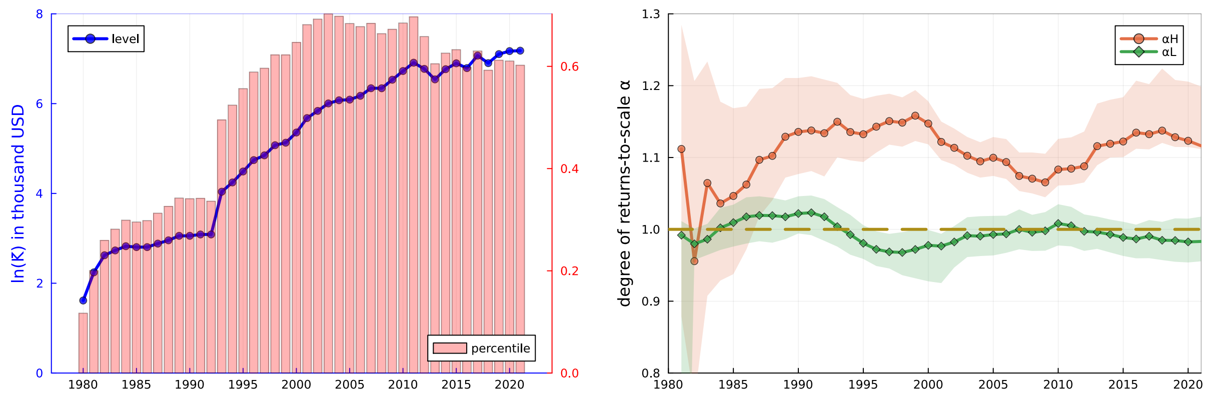
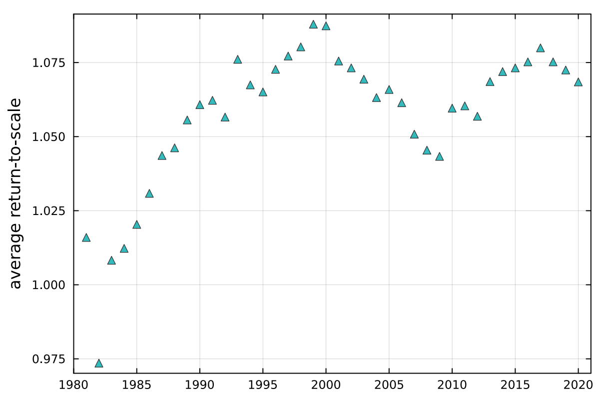


Figure 3: **Long-run Changes in Corporate Production Function: Alternative Indicators**

Notes: Graph (a) presents the percentile position of the estimated threshold over time. For each year t, we compute the relative rank of the convexity-concavity threshold as $\bar{k}_{pc} = \hat{F}_k(\hat{\bar{k}}_t)$, where $\hat{F}_k$ is the empirical cumulative distribution function of total capital, and $\bar{k}_t$ is the cross-sectional convexity-concavity threshold using Bayesian Markov Chain Monte Carlo (MCMC) estimation method. Graph (b) presents an average measure of returns-to-scale, calculated as $\alpha^H \times \bar{k}_{pc} + \alpha^L \times (1 - \bar{k}_{pc})$. Graph (c) provides a measure of the relative importance of convexity versus concavity in production, calculated as the ratio $\frac{\text{convexity}}{\text{concavity}} = \frac{\alpha^H \times \bar{k}_{pc}}{\alpha^L \times (1-\bar{k}_{pc})}$. Total output y is measured as Compustat data item *SALE* and capital stock k is the sum of physical capital (Compustat data item *PPENT*) and intangible capital, where we measure the stock of intangible capital by following Peters and Taylor (2017). Data is obtained from *Compustat*.

(a) average return-to-scale



(b) relative importance

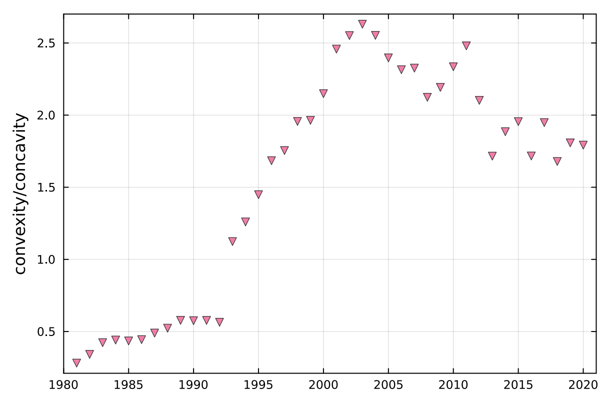


Figure 4: **Trends in Average Return-to-scale Across Industries**

Notes: This figure presents an average measure of returns-to-scale for different industries, calculated as $\alpha^H \times \bar{k}_{pc} + \alpha^L \times (1 - \bar{k}_{pc})$. For each year t and each industry, we compute the relative rank of the convexity-concavity threshold as $\bar{k}_{pc} = \hat{F}_k(\hat{\bar{k}}_t)$, where $\hat{F}_k$ is the empirical cumulative distribution function of total capital, and $\bar{k}_t$ is the cross-sectional convexity-concavity threshold using Bayesian Markov Chain Monte Carlo (MCMC) estimation method. The industry classification follows Fama-French 10-industry approach. Total output y is measured as Compustat data item *SALE* and capital stock k is the sum of physical capital (Compustat data item *PPENT*) and intangible capital, where we measure the stock of intangible capital by following Peters and Taylor (2017). Data is obtained from *Compustat*.

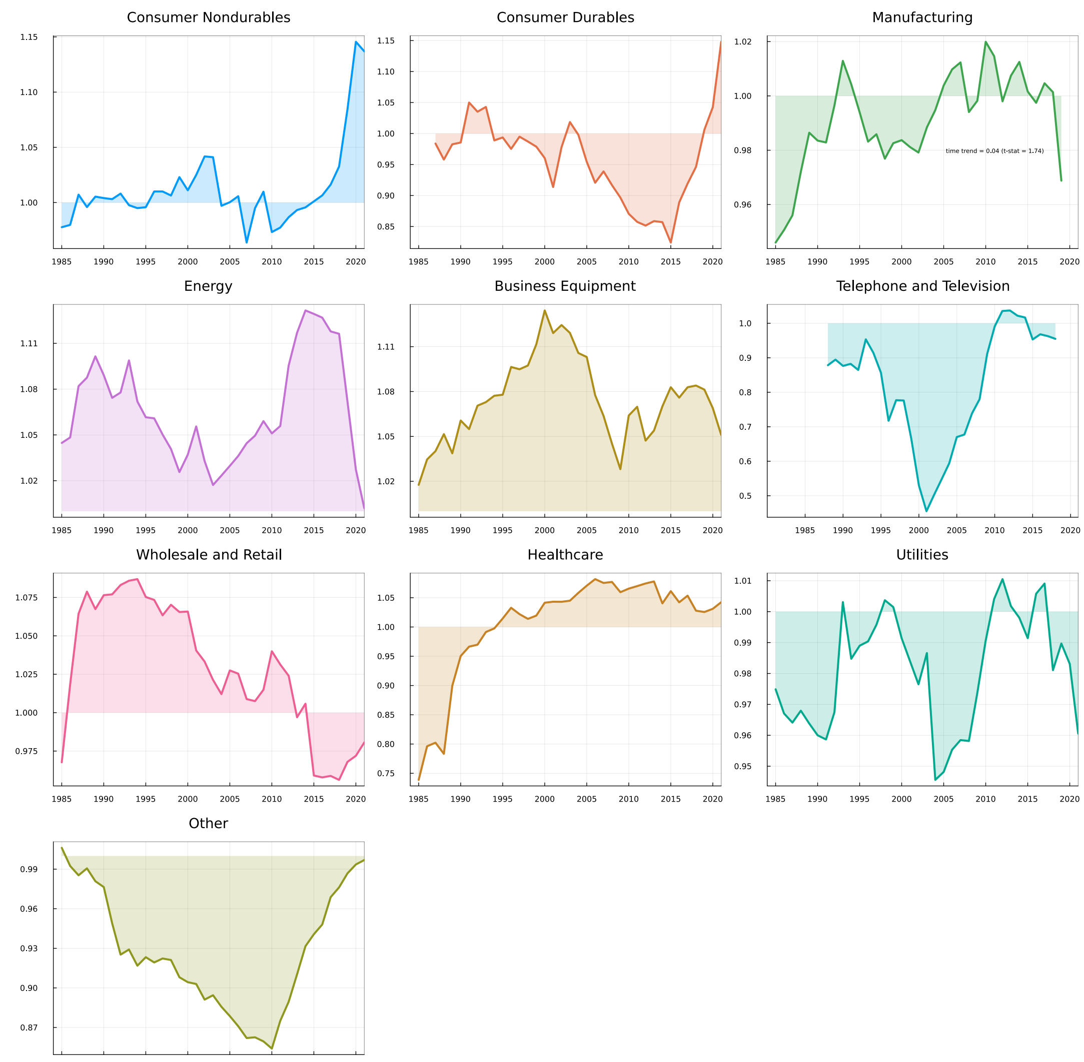


Figure 5: **Trends in Average Return-to-scale Across Countries**

Notes: This figure presents an average measure of returns-to-scale for different countries, calculated as $\alpha^H \times \bar{k}_{pc} + \alpha^L \times (1 - \bar{k}_{pc})$. For each year t and each country, we compute the relative rank of the convexity-concavity threshold as $\bar{k}_{pc} = \hat{F}_k(\hat{\bar{k}}_t)$, where $\hat{F}_k$ is the empirical cumulative distribution function of total capital, and $\bar{k}_t$ is the cross-sectional convexity-concavity threshold using Bayesian Markov Chain Monte Carlo (MCMC) estimation method. Total output y is measured as Compustat data item *SALE* and capital stock k is the sum of physical capital (Compustat data item *PPENT*) and intangible capital, where we construct the stock of intangible capital by following Peters and Taylor (2017). Data is obtained from *Compustat Global*.

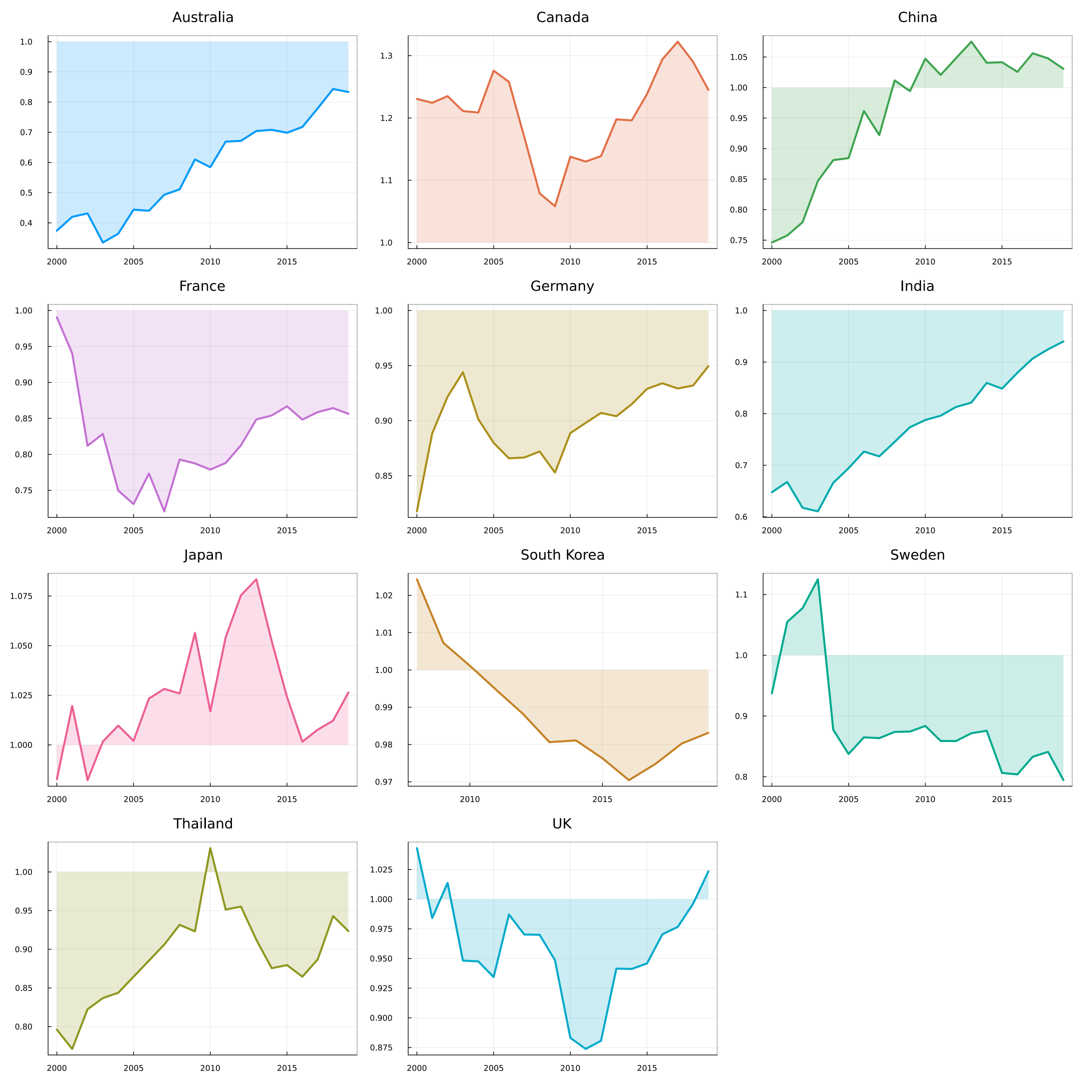


Figure 6: **Binscatter Plot between Convexity-Concavity Threshold and Share of Firms with Negative Net Earnings**

Notes: This figure presents the binscatter plot between the industry-level turning point for the convexity-concavity production function and the share of firms with negative earnings. The red dashed line represents the linear-fit regression. Specifically, for each year and each industry, we obtain the empirical measures of turning point with our baseline approach. In addition, for each industry in each year, we count the number of firms with negative profits and divide it by the total number of firms to obtain the industry-level share of firms with negative net earnings. Data is obtained from *Compustat*.

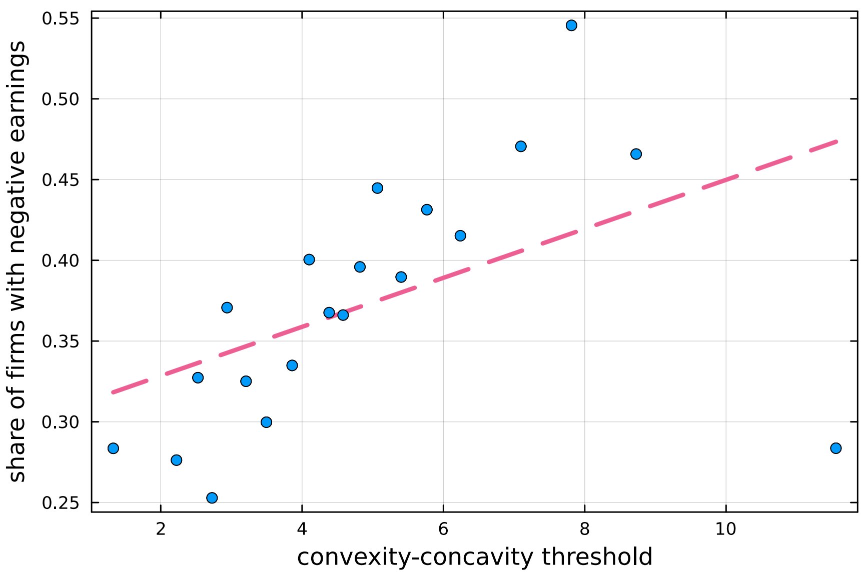
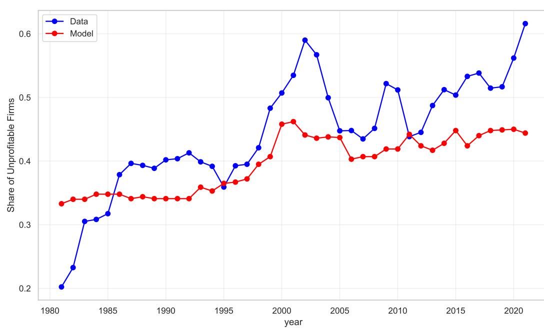


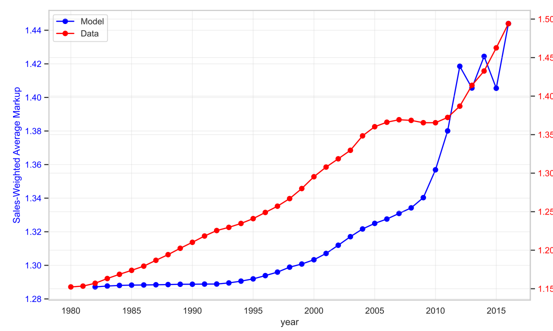
Figure 7: **Quantitative Exercises: Data v.s. Model**

Notes: This figure illustrates the time series of the share of unprofitable firms, markup, and markup distributions, as calculated from both the empirical data and the model. For each year, we determine the proportion of firms with negative net earnings by dividing the number of firms reporting losses by the total number of firms. Firm-level markups are estimated following the methodology of De Loecker, Eeckhout and Unger (2020), after which we compute the sales-weighted average with 10-year rolling window to be consistent with our baseline threshold measure.

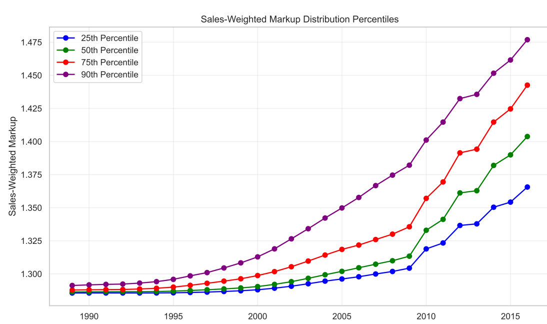
(a) share of unprofitable firms



(b) markup



(c) markup percentiles



(d) markup percentiles by De Loecker, Eeckhout and Unger (2020)

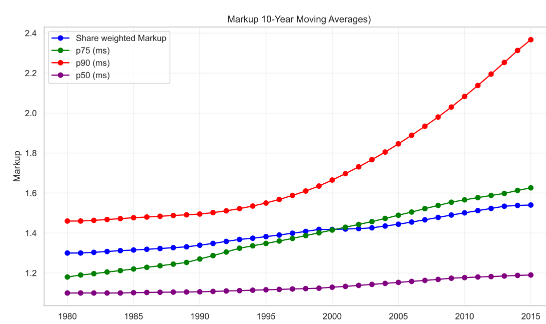
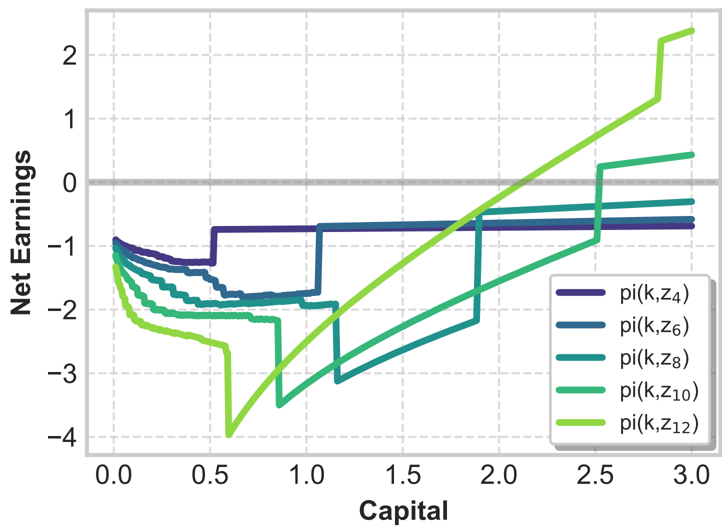


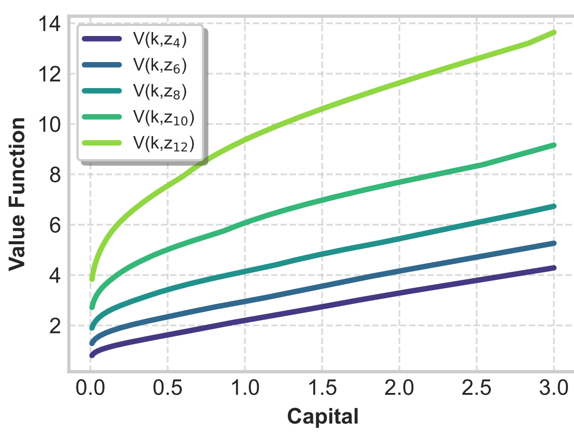
Figure 8: **Firm Heterogeneity, Investment Decisions, and Net Earnings**

Notes: This figure depicts the value function, investment decisions, and net earnings for firms across different levels of productivity and capital stock. The subscript of productivity z increases with the firm's productivity level. Details on the model setup and parameter choices can be found in Section D of the appendix.

(a) Net Earnings



(b) Value Function



(c) Investment Decision

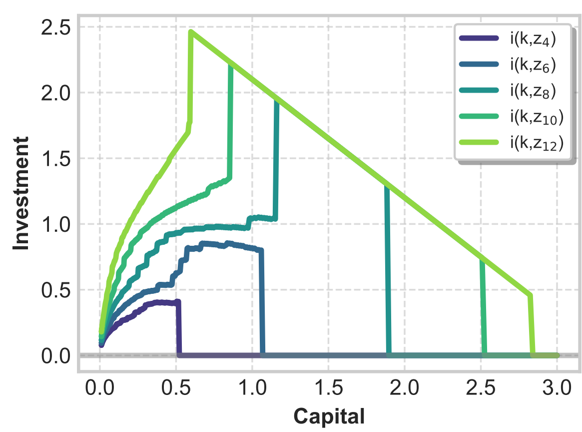


Table 1: **Reduced-form Evidence: Convexity-Concavity Threshold and Share of Firms with Negative Net Earnings**

Notes: This table presents the association between industry-level convexity-concavity threshold and the share of firms with negative earnings with different fixed-effect model specifications. Specifically, for each year and each industry, we obtain the empirical measures of turning point with our baseline approach. In addition, for each industry in each year, we count the number of firms with negative profits and divide it by the total number of firms to obtain the industry-level share of firms with negative net earnings. Original data used in this table is at the industry-year level and obtained from *Compustat*. T-statistics are in parentheses. *, **, and *** represent results significant at the 10%, 5%, and 1% levels, respectively. Standard errors are clustered at the industry level.

	share of firms with negative earnings			
	(1)	(2)	(3)	(4)
Threshold	0.012***	0.007*	0.019***	0.010***
	(3.674)	(1.786)	(7.388)	(3.525)
Year		Yes		Yes
Industry			Yes	Yes
N	420	420	420	420
Adjusted R^2	0.029	0.070	0.695	0.799

Table 2: **Overview of Time Trends Across Baseline, Industries, and Countries**

Notes: This table summarizes the time trend coefficient for our baseline analysis, industry-level evidence, and global investigation. Threshold $\bar{k}$ is the cross-sectional convexity-concavity threshold using Bayesian Markov Chain Monte Carlo (MCMC) estimation method. Threshold in Percentile $\bar{k}_{pc}$ presents the percentile position of the estimated threshold over time. For each year t, we compute the relative rank of the convexity-concavity threshold as $\bar{k}_{pc} = \hat{F}_k(\hat{\bar{k}}_t)$, where $\hat{F}_k$ is the empirical cumulative distribution function of total capital. Average returns-to-scale is calculated as $\alpha^H \times \bar{k}_{pc} + \alpha^L \times (1 - \bar{k}_{pc})$. The relative importance of convexity versus concavity is calculated as the ratio $\frac{\text{convexity}}{\text{concavity}} = \frac{\alpha^H \times \bar{k}_{pc}}{\alpha^L \times (1-\bar{k}_{pc})}$. α^H and α^L are the estimated degrees of return-to-scale before and after the threshold. Total output y is measured as Compustat data item *SALE* and capital stock k is the sum of physical capital (Compustat data item *PPENT*) and intangible capital, where we measure the stock of intangible capital by following Peters and Taylor (2017). Data is obtained from *Compustat*. T-statistics are in parentheses. *, **, and *** represent results significant at the 10%, 5%, and 1% levels, respectively.

Time-Trend Estimates		**Threshold**	**Threshold In Percentile**	**Average Return-to-scale**	**Convexity/Concavity**
Aggregate Evidence	**Baseline**	0.146***	0.012***	0.001***	0.051***
		(25.506)	(8.707)	(6.066)	(7.729)
Industry Evidence	**Consumer Nondurables**	0.055***	0.0016	0.0016**	0.0050*
		(6.30)	(1.56)	(2.93)	(1.86)
	Consumer Durables	0.097***	0.0058***	-0.0021**	0.0073*
		(17.87)	(5.80)	(-2.02)	(1.84)
	Manufacturing	0.10***	0.0046***	0.0052*	0.0083***
		(26.15)	(7.45)	(1.72)	(5.57)
	Energy	0.052***	-0.0081***	0.00035	-0.088***
		(4.47)	(-6.08)	(0.66)	(-3.59)
	Business Equipment	0.11***	0.0064***	0.00039	0.022***
		(17.73)	(4.72)	(0.90)	(2.86)
	Telephone and Television	0.066***	-0.0032	0.0055*	-0.025***
		(3.63)	(-1.31)	(1.81)	(-2.56)
	Wholesale and Retail	0.057***	-0.00040	-0.004***	-0.0036**
		(13.14)	(-1.31)	(-6.50)	(-3.26)
	Healthcare	0.13***	0.0075***	0.0059***	0.24***
		(6.24)	(8.66)	(7.46)	(7.13)
	Utilities	0.094***	0.0013***	0.00052*	0.0027***
		(17.05)	(4.86)	(1.87)	(6.25)
	Other	0.14***	0.00043***	-0.00075	0.66***
		(32.57)	(9.99)	(-1.07)	(9.37)
Global Evidence	**Australia**	0.047	0.002	0.027***	0.130***
		(1.394)	(0.687)	(16.170)	(4.656)
	Canada	0.096***	0.015***	0.002	-0.002
		(5.912)	(5.569)	(0.832)	(-0.062)
	China	0.040***	-0.005***	0.016***	0.006*
		(2.982)	(-3.468)	(8.114)	(1.915)
	France	0.013	0.016***	0.000	-0.054
		(0.457)	(3.967)	(0.137)	(-1.284)
	Germany	-0.002	0.010***	0.003**	-0.003
		(-0.233)	(6.201)	(3.036)	(-0.899)
	India	0.034*	0.002	0.017***	-0.004
		(1.696)	(0.907)	(21.208)	(-1.491)
	Japan	0.027	0.003***	0.002*	0.003
		(1.621)	(4.713)	(1.770)	(1.175)
	Korea	0.213***	0.002***	-0.004***	0.000
		(8.282)	(4.938)	(-4.755)	(-0.364)
	Sweden	-0.041*	0.012***	-0.011***	-0.006
		(-1.785)	(4.084)	(-4.379)	(-0.864)
	Thailand	0.000	-0.001	0.006**	0.001
		(0.008)	(-0.463)	(2.761)	(0.159)
	UK	-0.135***	-0.013**	-0.001	0.056
		(-3.084)	(-2.030)	(-0.640)	(1.480)

Online Appendix

(Not For Publication)

This version: June 17, 2025

A Implementation details of Estimating the Shape of Corporate Production Function

A.1 First-step Bayesian MCMC

Bayesian MCMC is a simulation-based sampling method that has gained significant traction in economic and financial research. MCMC offers an efficient framework for estimating continuous structural breaks, which are inherently more complex than discrete ones. For high-dimensional joint posterior distributions, the MCMC sampler iteratively samples from each univariate conditional posterior, progressively constructing the overall distribution. Under mild regularity conditions, these conditionally sampled densities converge to approximate the target joint posterior distribution.[1]

In particular, denote the observed data as $\mathcal{Y}_t = \{Y_{it}, i = 1, \ldots, N_t\}$ and $\mathcal{K}_t = \{K_{it}, i = 1, \ldots, N_t\}$. The joint likelihood $f(\mathcal{Y}_t, \mathcal{K}_t \mid \boldsymbol{\theta}_t)$ is proportional to

$$\frac{1}{\sqrt{2\pi\sigma_H^2}^{n_1}\sqrt{2\pi\sigma_L^2}^{n_2}} \exp\left\{-\frac{\sum_{i:K_{i,t}<\overline{K}_t}[\log Y_{i,t} - (a_t^H + \omega_{i,t} + \alpha_t^H \log K_{i,t})]^2}{2\sigma_H^2} - \frac{\sum_{i:K_{i,t}\geq\overline{K}_t}[\log Y_{i,t} - (a_t^L + \omega_{i,t} + \alpha_t^L \log K_{i,t})]^2}{2\sigma_L^2}\right\} \tag{A1}$$

where n_1 and n_2 are the number of observations with $K_{i,t} < \overline{K}_t$ and $K_{i,t} \geq \overline{K}_t$ for a given $\overline{K}_t$.

Bayesian statistics differ from frequentist approaches by treating parameters as random variables, where prior distributions capture the researchers' prior beliefs about these parameters. Denote the priors on parameters as $\pi(\boldsymbol{\theta})$. The posterior is then proportional to $f(\mathcal{Y}_t, \mathcal{K}_t \mid \boldsymbol{\theta}_t)\pi(\boldsymbol{\theta})$. Since MCMC draws from posterior distributions, interval estimates are readily available. For a point estimate $\hat{\theta}$, we approximate its *credible interval* as $[\hat{\theta} - 2\hat{s}, \hat{\theta} + 2\hat{s}]$, where $\hat{s}$ represents the estimated posterior standard deviation. For further details on MCMC in structural break applications, please refer to Bradley P Carlin, Alan E Gelfand and Adrian FM Smith (1992) and Stephens (1994). In our estimation, we tune the proposal density with up to 2,000 iterations and generate 20,000 Monte Carlo samples using a Metropolis-Hastings embedded Gibbs sampler.

A.2 Second-step TFP estimation

In production function estimation, endogeneity is a critical issue that can bias and skew total factor productivity (TFP) estimates. This problem arises because firms' input choices, such as labor, capital, and materials, are often responsive to unobserved productivity shocks. For example, a

[1] For an in-depth discussion of Bayesian MCMC methods, refer to Erica X.N. Li, Haitao Li, Shujing Wang and Cindy Yu (2019) and Erica X.N. Li, Guoliang Ma, Shujing Wang and Cindy Yu (2021).

firm experiencing a positive productivity shock (i.e., shock to ω) may increase input levels, creating simultaneity bias. Traditional ordinary least squares (OLS) cannot disentangle the effect of productivity shocks from the true contribution of inputs to output, often leading to inflated or deflated coefficients.

To address this endogeneity, Levinsohn and Petrin (2003) proposed a method that uses intermediate inputs, such as materials or energy, as proxies to control for unobserved productivity shocks. This approach assumes that intermediate inputs are more flexible and responsive to productivity changes than capital, which is typically fixed in the short term. By conditioning on intermediate inputs, the Levinsohn and Petrin method accounts for productivity shocks in input decisions, isolating the actual effect of capital and labor on output. This adjustment results in more accurate production function estimates, essential for reliable TFP measurement. We adopt this approach by using production costs (*Compustat* data item *COGS*) as a proxy for materials due to the unavailability of direct measures like electricity or fuel in our firm-level data. Additionally, we implement the Olley and Pakes (1996) approach as a robustness check, using our constructed total investment and firm age as control variables.

B Relationship among Return-to-scale, Fixed Cost and Marginal Cost

B.1 Cost structure and return-to-scale

Here we use an example to formally establish this point. By definition, a production function F exhibits increasing returns to scale if $F(\lambda K, \lambda L) > \lambda F(K, L), \quad$ for any $\lambda > 1$. In cost terms, this condition implies that total cost satisfies $TC(\lambda q) < \lambda \cdot TC(q)$. Firms generally face two categories of costs. Fixed costs (FC) are incurred independently of output and typically represent one-time expenditures. Variable costs (VC), by contrast, depend on the quantity produced. Due to learning effects or scale economies, variable costs may decrease with output, leading to declining marginal costs. The total cost function can be expressed as $C(q) = FC + \int_0^q MC(z)\,dz$, where $MC(q)$ denotes the marginal cost of producing quantity q. For analytical convenience, assume that marginal cost follows the functional form $MC(q) = c \cdot q^{-\alpha}$, where $c > 0$ and $0 < \alpha < 1$. The parameter α governs the rate of cost decline: the larger α, the faster marginal cost decreases with output.

To obtain total cost, we first compute variable cost by integrating the marginal cost function $VC(q) = \int_0^q c \cdot z^{-\alpha} dz = \frac{c}{1-\alpha} q^{1-\alpha}$. Adding fixed cost yields $TC(q) = FC + \frac{c}{1-\alpha} q^{1-\alpha}$. Dividing total cost by output gives average cost $AC(q) = \frac{TC(q)}{q} = \frac{FC}{q} + \frac{c}{1-\alpha} q^{-\alpha}$. Both components decline with output: the fixed cost term shrinks as it is distributed over more units, and the variable cost term falls due to diminishing marginal cost. Hence, average cost decreases with output, reflecting increasing returns to scale. To verify this formally, consider the scaled cost expression: $\lambda \cdot TC(q) = \lambda FC + \lambda \cdot \frac{c}{1-\alpha} q^{1-\alpha}$. In comparison, we compute $TC(\lambda q) = FC + \frac{c}{1-\alpha}(\lambda q)^{1-\alpha} = FC + \frac{c}{1-\alpha}\lambda^{1-\alpha} q^{1-\alpha}$. Since $\lambda^{1-\alpha} < \lambda$ for $\lambda > 1$ and $0 < \alpha < 1$, the second term in $TC(\lambda q)$ grows less than proportionally in λ. If FC is sufficiently large, this sublinear growth becomes more pronounced, so the inequality $TC(\lambda q) < \lambda \cdot TC(q)$ holds, confirming the presence of increasing returns to scale.

B.2 Cost structure and convex-concave threshold

Consider a convex-concave production function $F(x)$ that features increasing marginal returns in the initial phase and decreasing marginal returns beyond a certain threshold. Specifically, the function exhibits a convex region (increasing returns to scale) when $x \leq x^*$ and a concave region (decreasing returns to scale) when $x > x^*$. In the convex region, production benefits from scaling advantages, such as those present in a startup phase. In the concave region, production becomes less efficient, possibly due to congestion or coordination issues at high input levels.

Mathematically, the production function is defined as $F(x) = \begin{cases} x^\alpha, & \text{if } x \leq x^*, \quad \alpha > 1 \\ x^* \cdot \left(\frac{x}{x^*}\right)^\beta, & \text{if } x > x^*, \quad \beta < 1 \end{cases}$. The threshold x^* represents the input level at which the curvature of the production function changes. The corresponding output threshold is $y^* = F(x^*) = (x^*)^\alpha$.

The firm's original cost structure includes a fixed cost FC, which is independent of output, and a variable cost VC that reflects the cost of inputs required to produce a given output y. Assuming an input price of $w = 1$ for simplicity, the variable cost can be expressed as $VC(y) = F^{-1}(y)$, where $F^{-1}(y)$ is the inverse of the production function.

For outputs less than or equal to y^*, the inverse function takes the form $F^{-1}(y) = y^{1/\alpha}$, implying that $VC(y) = y^{1/\alpha}$. For outputs exceeding y^*, the inverse is $F^{-1}(y) = x^* \cdot \left(\frac{y}{y^*}\right)^{1/\beta}$, leading to a variable cost of $VC(y) = x^* \cdot \left(\frac{y}{(x^*)^\alpha}\right)^{1/\beta}$.

Marginal cost is the derivative of variable cost with respect to output: $MC(y) = \frac{d}{dy} VC(y)$. When $y \leq y^*$, this derivative becomes $MC(y) = \frac{1}{\alpha} y^{1/\alpha - 1}$. Since $\alpha > 1$, marginal cost decreases with output, reflecting economies of scale.

When $y > y^*$, marginal cost becomes $MC(y) = \frac{x^*}{\beta (x^*)^\alpha} y^{1/\beta - 1}$. Because $\beta < 1$, marginal cost increases with output, indicating diseconomies of scale.

Suppose technological improvements reduce variable costs by a factor k, where $0 < k < 1$. The new variable cost becomes $VC_{\text{new}}(y) = k \cdot VC_{\text{old}}(y)$. As a result, the new marginal cost is scaled accordingly: $MC_{\text{new}}(y) = k \cdot MC_{\text{old}}(y)$. For $y \leq y^*$, the new marginal cost becomes $MC_{\text{new}}(y) = \frac{k}{\alpha} y^{1/\alpha - 1}$, and for $y > y^*$, $MC_{\text{new}}(y) = \frac{kx^*}{\beta (x^*)^\alpha} y^{1/\beta - 1}$.

The threshold y^* marks the point where marginal cost transitions from decreasing to increasing. In the original setup, this threshold is given by $y^* = (x^*)^\alpha$. After the cost reduction, the new economic threshold y^*_{new} is determined by the intersection of the new marginal cost curves.

Although the technical threshold x^* remains unchanged (since k scales both marginal cost regions proportionally), the economic threshold shifts. Because lower marginal cost makes it more attractive to produce larger output before encountering diseconomies of scale, firms will tend to operate at a higher output level before entering the concave region of the production function.

An increase in fixed costs has no effect on the technical threshold x^*, as it does not affect the production function's curvature. However, there is an indirect effect: higher fixed costs encourage firms to produce at larger scales in order to spread those costs over more output. Nonetheless, the switch point x^* itself remains technically unchanged.

When both fixed costs increase and variable costs decrease, the effects combine in the following way. A reduction in variable costs delays the onset of rising marginal cost, effectively extending the convex region (increasing returns to scale) in economic terms. Simultaneously, higher fixed costs incentivize firms to scale up production, although the technical threshold x^* will only shift if the production function itself changes—that is, if the parameters α or β are altered due to innovation.

C Model Proof

C.1 Proof of Production Function Shape Representation

Proof. The firm's optimization problem is shown as follows:

$$\begin{aligned}
\tilde{P}_t \tilde{Y}_t - W_t L_t &= \left(\frac{1}{\tilde{H}_t}\right)^{-\frac{1}{\varepsilon_t}} \left(\tilde{Y}_t\right)^{1-\frac{1}{\varepsilon_t}} - W_t L_t && \text{(A2)} \\
&= \left(\frac{1}{\tilde{H}_t}\right)^{-\frac{1}{\varepsilon_t}} \left(\tilde{A}_t \left(K_t^{\gamma_t} L_t^{1-\gamma_t}\right)^{s_t}\right)^{1-\frac{1}{\varepsilon_t}} - W_t L_t && \text{(A3)}
\end{aligned}$$

Given the market wage w_t, the first order condition with respect to labor L_t gives:

$$\tilde{L}_t = \left[\tilde{H}_t \left[\left(1 - \frac{1}{\varepsilon_t}\right)(1-\gamma_t)s_t\right]^{\varepsilon_t} \left(\tilde{A}_t K_t^{\gamma_t s_t}\right)^{\varepsilon_t - 1}\right]^{\frac{1}{\varepsilon_t - (1-\gamma_t)s_t(\varepsilon_t - 1)}} \tag{A4}$$

In this way, we can rewrite the optimized sales $Y_t \equiv \tilde{P}_t \tilde{Y}_t$ as below:

$$
\begin{aligned}
Y_t \equiv \tilde{P}_t \tilde{Y}_t &= \left[\tilde{H}_t \left(1 - \frac{1}{\varepsilon_t}\right)(1-\gamma_t)s_t\right]^{\frac{(\varepsilon_t-1)(1-\gamma_t)s_t}{\varepsilon_t-(1-\gamma_t)s_t(\varepsilon_t-1)}} \tilde{A}_t^{\frac{\varepsilon_t-1}{\varepsilon_t-(1-\gamma_t)s_t(\varepsilon_t-1)}} K_t^{\frac{(\varepsilon_t-1)\gamma_t s_t}{\varepsilon_t-(1-\gamma_t)s_t(\varepsilon_t-1)}} \\
&\equiv \tilde{Z}_t^{1-\alpha_t} K_t^{\alpha_t}
\end{aligned}
\tag{A5}
$$

where

$$
\alpha_t \equiv \frac{\gamma_t s_t \left(1 - \frac{1}{\varepsilon_t}\right)}{1 - (1-\gamma_t)\, s_t \left(1 - \frac{1}{\varepsilon_t}\right)} \tag{A6}
$$

$$
\tilde{Z}_t \equiv \left[\tilde{H}_t \left(1 - \frac{1}{\varepsilon_t}\right)(1-\gamma_t)s_t\right]^{\frac{(1-\frac{1}{\varepsilon_t})(1-\gamma_t)s_t}{1-s_t(1-\frac{1}{\varepsilon_t})}} \tilde{A}_t^{\frac{1-\frac{1}{\varepsilon_t}}{1-s_t(1-\frac{1}{\varepsilon_t})}} \tag{A7}
$$

□

C.2 Proof of Lemma 1

Proof. We set up the current-value Hamiltonian function as below:

$$
\mathcal{H}(K, I, q) \equiv F(K) - I - G(I, K) + q(I - \delta K), \tag{A8}
$$

where q is the Lagrangian multiplier associated with capital stock dynamics. For each $t \geq 0$ we maximize $\mathcal{H}$ with respect to the control variable I, which gives us

$$
1 + G_I(I, K) = q
$$

Next, we partially differentiate $\mathcal{H}$ with respect to the state variable and set the result equal to $rq - \dot{q}$, where r is the discount rate:

$$
\frac{\partial H}{\partial K} = F_K(K) - G_K(I, K) - \delta q = rq - \dot{q}. \tag{A9}
$$

The standard infinite horizon transversality condition says that irrespective of the time path of the capital stock, optimality requires that the present value of the shadow price itself, when discounted by $r + \delta$, is asymptotically zero, i.e.,

$$
\lim_{t\to\infty} q_t e^{-\int_0^t (r_\tau + \delta) d\tau} = 0. \tag{A10}
$$

In this way, q_t can be interpreted as the shadow price (measured in current output units) of capital along the optimal path. Multiplying by $e^{-\int_0^t (r_\tau+\delta)d\tau}$ on both sides of the previous equation, we get the following:

$$q_t = \int_t^\infty [F_K(K_s) - G_K(I_s, K_s)] e^{-\int_t^s (r_\tau+\delta)d\tau} ds > 0. \tag{A11}$$

As the first order condition shows that $1 + G_I(I, K) = q$, then we can show that the optimal investment I is an implicit function of the shadow price of capital q and the state variable K_t, i.e., $I_t \equiv \mathcal{M}(q_t, K_t)$. In addition, we can easily show that $\mathcal{M}$ is a strictly increasing function of q:

$$\frac{\partial I_t}{\partial q_t} = \frac{1}{G_{II}(M(q_t, K_t), K_t)} > 0 \tag{A12}$$

As G is homogeneous of degree 1, we can rewrite $I_t \equiv \mathcal{M}(q_t, K_t) = m(q_t) K_t$, where again m is an increasing function of q.

With the definitions on $\mathcal{A}(K, L)$ and $\mathcal{B}(I, K)$, we can rewrite π as follows:

$$\begin{aligned}
\pi &= F(K) - G(I, K) - I \\
&= \mathcal{A}(K) + F_K K + \mathcal{B}(I, K) - G_I I - G_K K - I \\
&= \mathcal{A}(K) + \mathcal{B}(I, K) + (F_K K - G_K K) - (1 + G_I) I \\
&= \mathcal{A}(K) + \mathcal{B}(I, K) + (F_K K - G_K K) - (1 + G_I) m(q) K \\
&= \mathcal{A}(K) + \mathcal{B}(I, K) + (F_K K - G_K K) - (1 + G_I) m \left(\int_t^\infty e^{-\int_t^s r_\tau d\tau} [F_{K_s} K_s - G_{K_s} K_s] ds \right) K
\end{aligned} \tag{A13}$$

As G is homogeneous of degree 1, we have $\mathcal{B}(I, K) = 0$. If F is a convex function, then we have $\mathcal{A}(K) = F(K) - F_K(K)K < 0$. More importantly, as F is convex, then it means that $F_K < F_{K_s}$ if $K < K_s$. If the degree of convexity is sufficiently large and K is small such that G_I is large and G_K is small, then we have $(F_K K - G_K K) - (1 + G_I) m \left(\int_t^\infty e^{-\int_t^s r_\tau d\tau} [F_{K_s} K_s - G_{K_s} K_s] ds \right) K < 0$ as m is a strictly increasing function. In this case, the net earnings become negative, i.e., $\pi < 0$.

□

C.3 Proof of Lemma 2

Proof. The value of the firm as seen from time t is

$$V_t = \int_t^\infty (F(K_\tau) - G(I_\tau, K_\tau) - I_\tau) e^{-\int_t^\tau r_s ds} d\tau \tag{A14}$$

Consequently, when moving along the optimal path,

$$V_t = V^*(K_t, t) = \int_t^\infty \left(\mathcal{A}(K_\tau) + \mathcal{B}(I_\tau, K_\tau)\right) e^{-\int_t^\tau r_s ds} d\tau \tag{A15}$$

$$+ \int_t^\infty \left([F_K - G_K]K_\tau - (1 + G_I)I_\tau\right) e^{-\int_t^\tau r_s ds} d\tau$$

$$= \int_t^\infty \left(\mathcal{A}(K_\tau) + \mathcal{B}(I_\tau, K_\tau)\right) e^{-\int_t^\tau r_s ds} d\tau + q_t K_t$$

Isolating q_t, it follows that

$$q_t^m \equiv q_t = \frac{V_t}{K_t} - \frac{1}{K_t}\int_t^\infty \left(\mathcal{A}(K_\tau) + \mathcal{B}(I_\tau, K_\tau)\right) e^{-\int_t^\tau r_s ds} d\tau \tag{A16}$$

$\square$

D Quantitative Model Setup

Time is discrete and is indexed by $t = 1, 2, \ldots$. The horizon is infinite. At time t, a positive mass of price-taking firms produce a homogeneous good by means of the production function $y_{i,t} = z_{i,t}\left(k_{i,t}^{\beta} l_{i,t}^{1-\beta}\right)^{\alpha_t^H \times \mathbb{I}_{k_{i,t} < \bar{k}_t} + \alpha_t^L \times \mathbb{I}_{k_{i,t} \leq \bar{k}_t}}$. With $k_{i,t}$ we denote total capital, $l_{i,t}$ as labor, β is the capital share, and z_t is the idiosyncratic random disturbance. As discussed

The dynamics of the idiosyncratic component z_t is described by

$$\log z_{t+1} = \rho_z \log z_t + \sigma_z \varepsilon_{z,t+1} \tag{A17}$$

with $\varepsilon_{z,t} \sim N(0,1)$ for all $t \geq 0$. The conditional distribution of z_t will be denoted as $H(z_{t+1}|z_t)$.

Firms discount future profits by means of the time-invariant factor $\frac{1}{R}$, $R > 1$. Adjusting the capital stock by x bears a cost $g(x, k)$. Capital depreciates at the rate $\delta \in (0, 1)$.

We assume that the demand for the firm's output and the supply of capital are infinitely elastic and normalize their prices at 1. Operating firms incur a cost $c_f > 0$, drawn from the common time-invariant distribution G. Firms that quit producing cannot re-enter the market at a later stage and recoup the undepreciated portion of their capital stock, net of the adjustment cost of driving it to 0.

Every period there is a constant mass $M > 0$ of prospective entrants, each of which receives a signal q about her productivity, with $q \sim Q(q)$. Conditional on entry, the distribution of the idiosyncratic shock in the first period of operation is $H(z'|q)$, strictly decreasing in q. Entrepreneurs that decide to enter the industry pay an entry cost $c_e \geq 0$.

The Incumbent's Optimization Program Upon exit, a firm obtains a value equal to the undepreciated portion of its capital k, net of the adjustment cost it incurs in order to dismantle it, i.e. $V_x(k) = k(1-\delta) - g\left[-k(1-\delta), k\right]$.

Then, the start-of-period value of an incumbent firm is given by the function $V(z,k)$ which solves the following functional equation:

$$V(z,k) = \pi(z,k) + \int \max\left\{V_x(k), \tilde{V}(z,k) - c_f\right\} dG(c_f) \tag{A18}$$

where

$$\tilde{V}(z,k) = \max_x -x - g(x,k) + \frac{1}{R}\int V(z',k')dH(z'|z) \tag{A19}$$

subject to $k' = k(1-\delta) + x$.

Entry For an aggregate state λ, the value of a prospective entrant that obtains a signal q is

$$V_e(q) = \max_{k'} -k' + \frac{1}{R}\int V(z',k')dH(z'|q) \tag{A20}$$

She will invest and start operating if and only if $V_e(q) \geq c_e$.

Recursive Competitive Equilibrium For given initial condition Γ_0, a recursive competitive equilibrium consists of (i) value functions $V(z,k)$, $\tilde{V}(z,k)$ and $V_e(q)$, (ii) policy functions $x(z,k)$, $k'(q)$, and (iii) bounded sequences of incumbents' measures $\{\Gamma_t\}_{t=1}^{\infty}$, and entrants' measures $\{E_t\}_{t=0}^{\infty}$ such that, for all $t \geq 0$,

1. $V(z,k)$, $\tilde{V}(z,k)$, and $x(z,k)$ solve the incumbent's problem;
2. $V_e(q)$ and $k'(q)$ solve the entrant's problem;
3. Capital market clears;
4. The entrant measure E_{t+1} and incumbent measure Γ_{t+1} satisfy their respective laws of motion

Functional Forms Investment adjustment costs are modeled as a sum of a fixed portion and a convex portion:

$$g(x,k) = \chi(x)c_0 k + c_1\left(\frac{x}{k}\right)^2 k, \quad c_0, c_1 \geq 0, \tag{A21}$$

where $\chi(x) = 0$ for $x = 0$ and $\chi(x) = 1$ otherwise.

Calibration The parameter values used in our quantitative exercise are shown in the following table.

Table A1: **Calibration**

Parameter	**Description**	**Value**
$\frac{1}{R}$	annual discount factor	0.9615
δ	annual depreciation rate	0.10
β	capital share	0.36
c_f	fixed cost	0.75
c_e	entry cost	1.5
c_1	convex adjustment cost parameter	0.25
c_0	non-convex/fixed adjustment cost	0.15
ρ_z	autocorrelation coefficient	0.62
σ_z	std. dev. of shocks	0.42

In our model, the markup here is computed exactly as in De Loecker, Eeckhout and Unger (2020), which is the relative ratio of price per unit to marginal production cost.

E Additional Figures and Tables

Figure A1: **The Rise of Firms with Negative Net Earnings: Ten Different Industries**

Notes: This figure presents the time-series plot of the fraction of unprofitable public firms in different industries. In each year, for each industry, we count the number of firms with negative net earnings and divide it by the total number of firms. Ten industries are defined as follows: *Consumer Nondurables* (SIC 0100-0999, 2000-2399, 2700-2749, 2770-2799, 3100-3199, 3940-3989); *Consumer Durables* (SIC 2500-2519, 2590-2599, 3630-3659, 3710-3711, 3714-3714, 3716-3716, 3750-3751, 3792-3792, 3900-3939, 3990-3999); *Manufacturing* (SIC 2520-2589, 2600-2699, 2750-2769, 2800-2829, 2840-2899, 3000-3099, 3200-3569, 3580-3621, 3623-3629, 3700-3709, 3712-3713, 3715-3715, 3717-3749, 3752-3791, 3793-3799, 3860-3899); *Oil, Gas, and Coal Extraction and Products* (SIC 1200-1399, 2900-2999); *Business Equipment* (SIC 3570-3579, 3622-3622, 3660-3692, 3694-3699, 3810-3839, 7370-7372, 7373-7373, 7374-7374, 7375-7375, 7376-7376, 7377-7377, 7378-7378, 7379-7379, 7391-7391, 8730-8734); *Telephone and Television Transmission* (SIC 4800-4899); *Wholesale, Retail, and Some Services* (SIC 5000-5999, 7200-7299, 7600-7699); *Healthcare, Medical Equipment, and Drugs* (SIC 2830-2839, 3693-3693, 3840-3859, 8000-8099); *Utilities* (SIC 4900-4949); and *Others*. Data is obtained from *Compustat*.

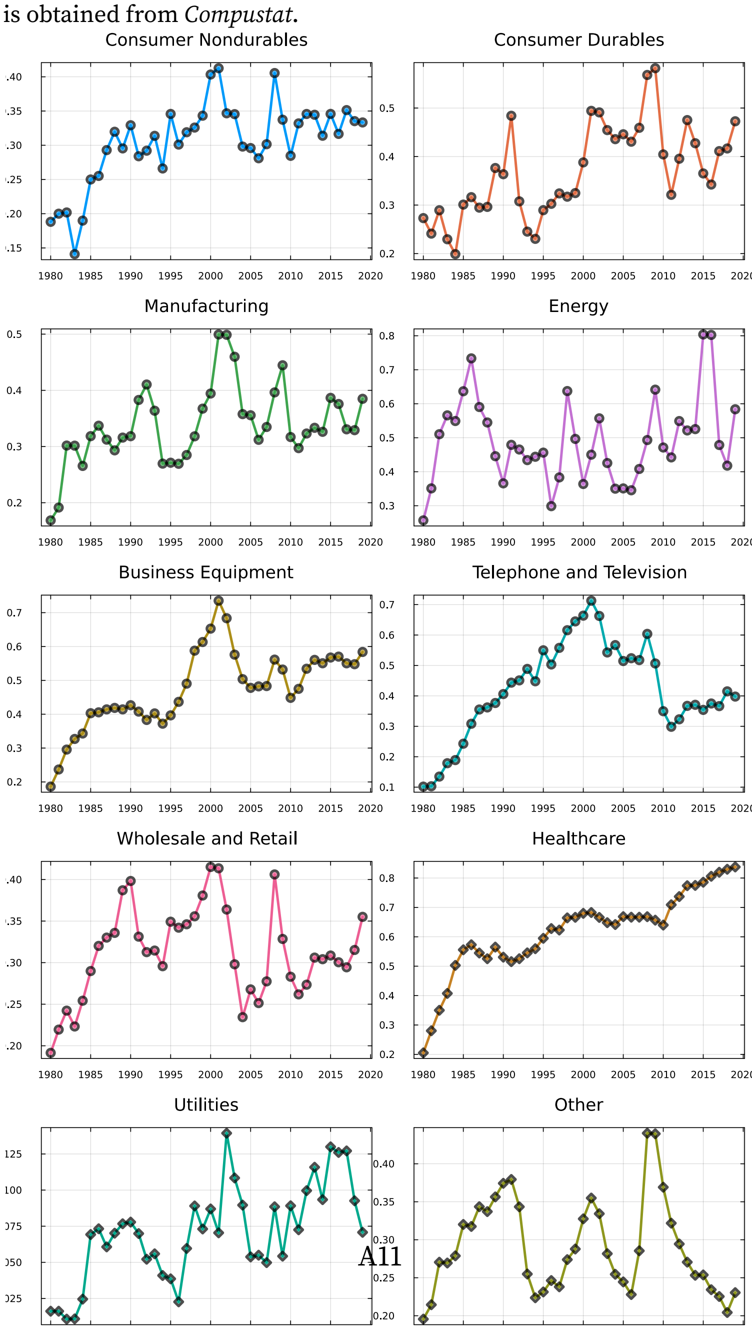


Figure A2: **Average Firm Age**

Notes: This figure presents the time-series plot of the average age for public companies in the US. A firm's age is defined as the year difference between the current year and the year that a certain firm first appears in the *Compustat* dataset. The fraction of young firms is computed as follows. For each year, we count the number of firms with an age less than 5 and divide it by the total number of firms. Data is obtained from *Compustat*.

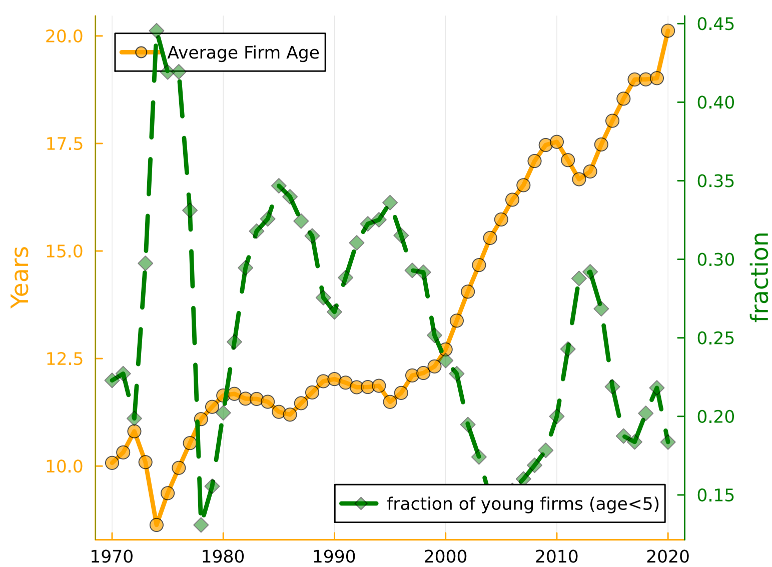


Figure A3: **The Rise of Firms with Negative Net Earnings: Different Stock Exchanges**

Notes: This figure presents the time-series plot of the fraction of unprofitable public firms in different stock exchanges. In each year, for each stock exchange, we count the number of firms with negative net earnings and divide it by the total number of firms in that exchange. Data is obtained from *Compustat*.

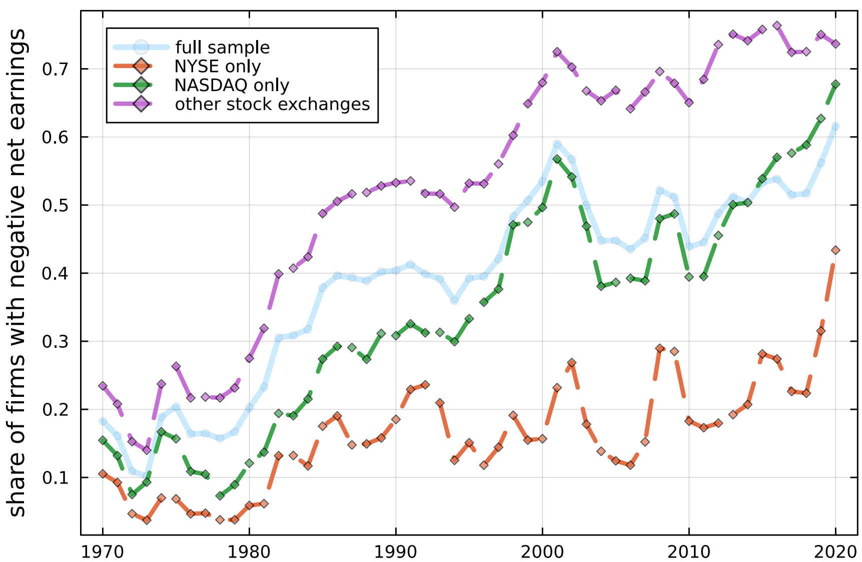


Figure A4: Potential Issues with Public Firms Only

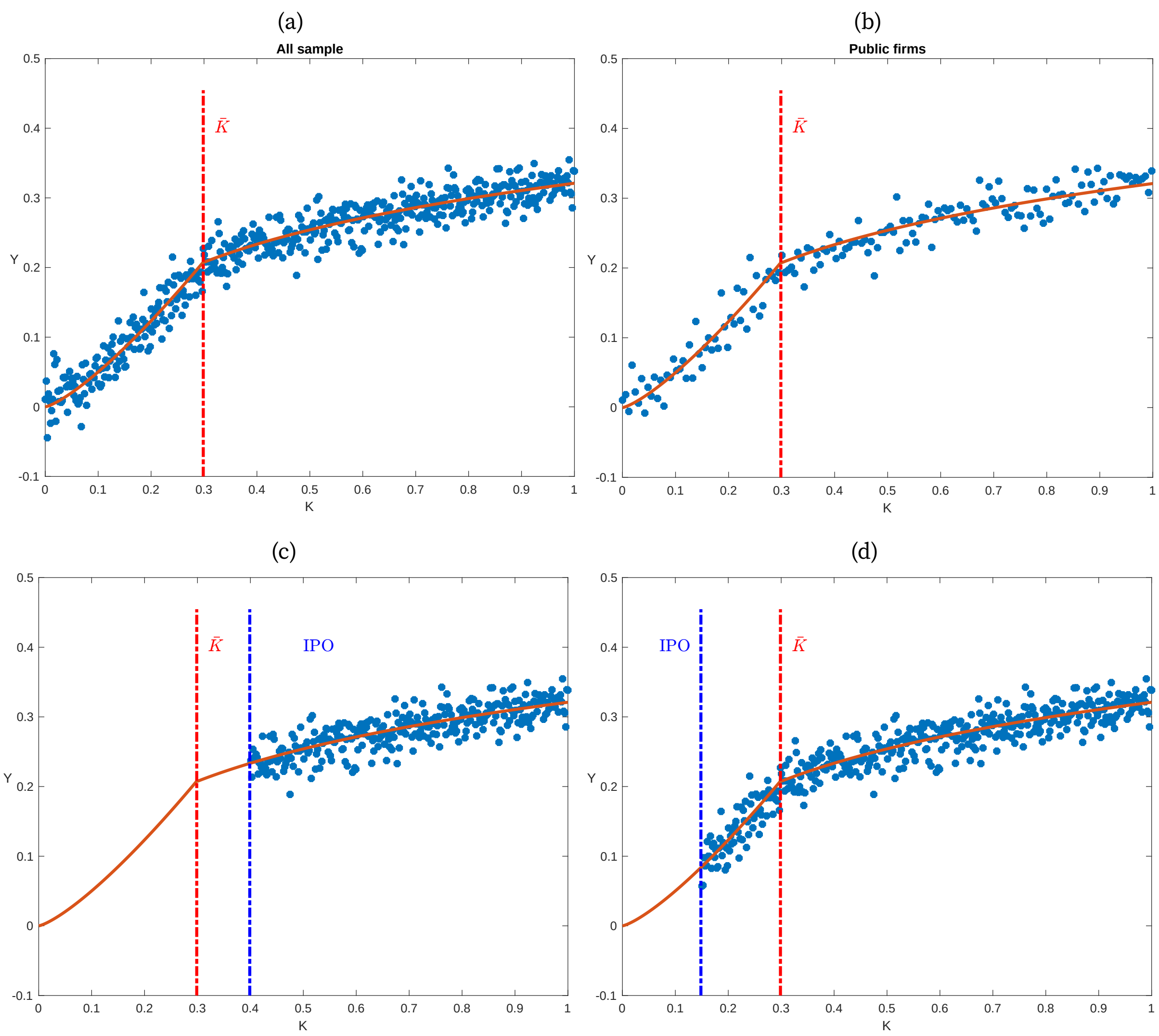
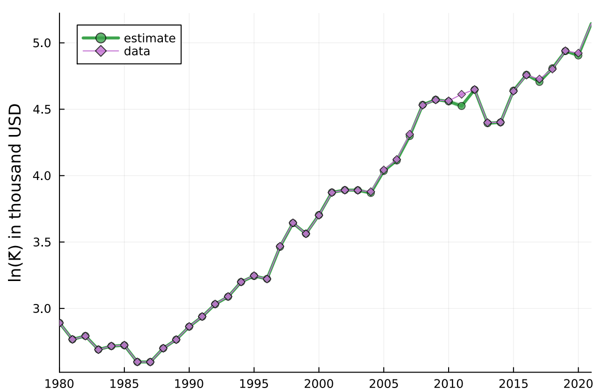


Figure A5: **Long-run Changes in Corporate Production Function: Simulation Exercise**

Notes: Graph (a) compares the simulated true and estimated structural breaks $\bar{k}_t$ obtained from Bayesian MCMC. The blue curve represents the true value, while the red represents the estimation. Graph (b) shows the corresponding estimated slope coefficients, and the green and orange dashed lines indicate the true values: 1.3 and 0.7. The blue curve represents the estimated time series $\hat{\beta}_{at}$ and the red curve represents the estimated time series $\hat{\beta}_{et}$. Total output y is measured as Compustat data item *SALE* and capital stock k is the sum of physical capital (Compustat data item *PPENT*) and intangible capital, where we measure the stock of intangible capital by following Peters and Taylor (2017). Data is obtained from *Compustat*.

(a) simulation: changepoint



(b) simulation: slope coefficient

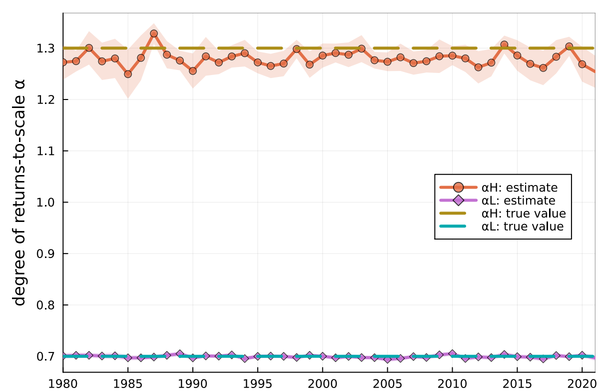
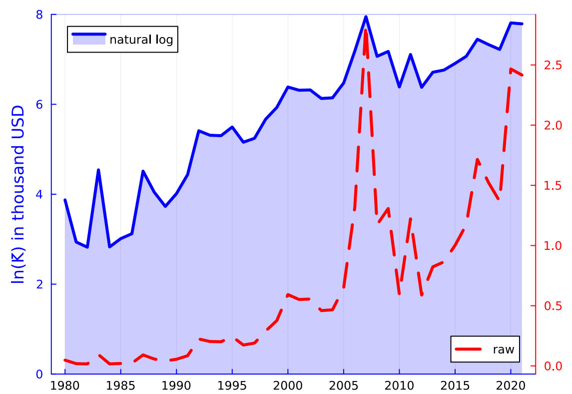


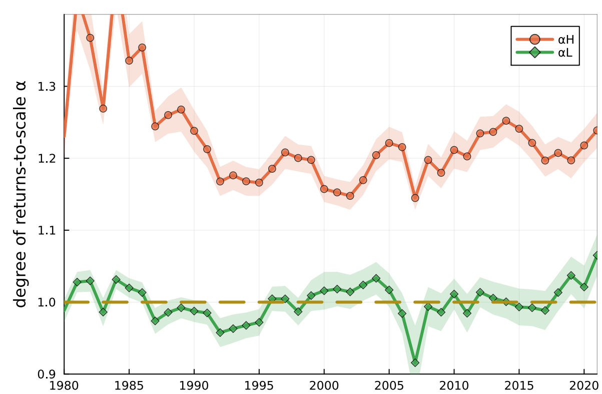
Figure A6: **Long-run Changes in Corporate Production Function: Alternative Functional Forms**

Notes: Graphs (a) and (b) report the estimated structural breaks and slope coefficients for continuous production function shown as in Equation (13). Meanwhile, Graph (c) and (d) report the corresponding results for using the exponential functional form shown as in Equation (14). Across all these model specifications, total output y is measured as Compustat data item *SALE* and capital stock k is the sum of physical capital (Compustat data item *PPENT*) and intangible capital, where we measure the stock of intangible capital by following Peters and Taylor (2017). Data is obtained from *Compustat*.

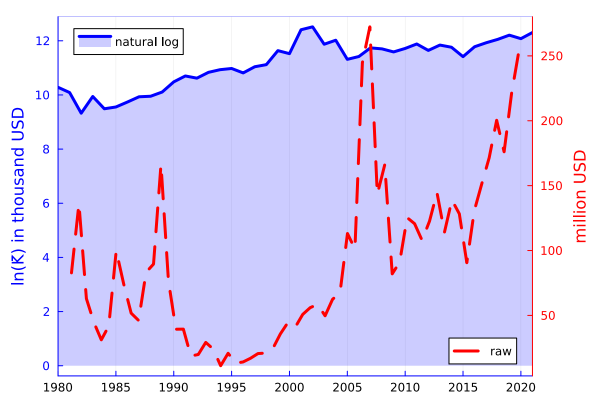
(a) continuous function: changepoint



(b) continuous function: slope coefficient



(c) exponential function: changepoint



(d) exponential function: coefficient

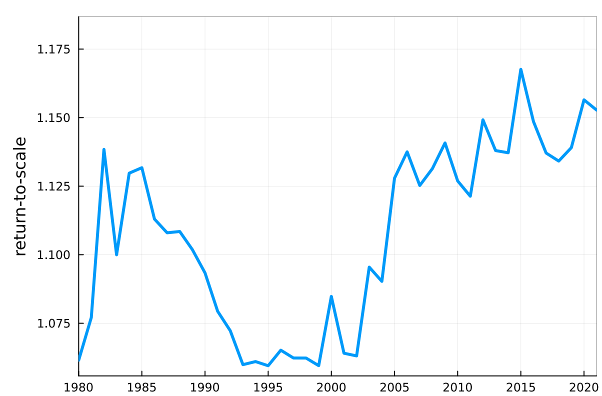
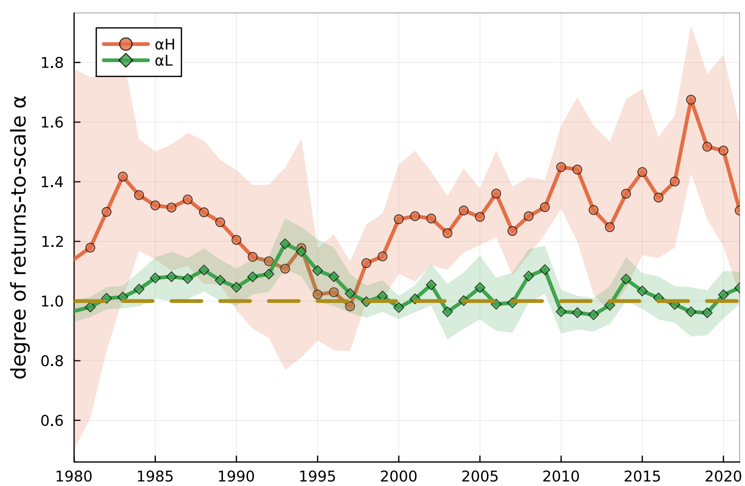


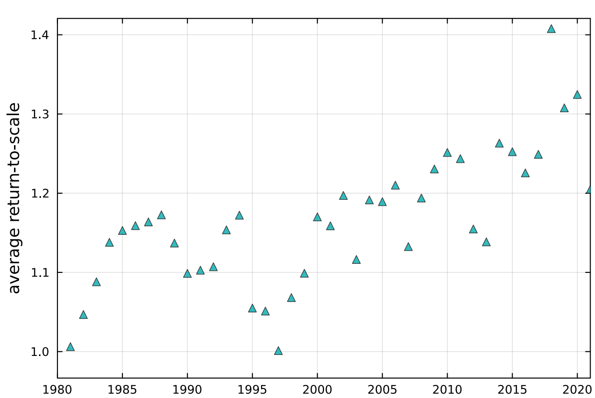
Figure A7: **Long-run Changes in Corporate Production Function with Olley and Pakes (1996) adjustment**

Notes: Graph (a) shows the estimated degree of return-to-scale, and the orange and green lines indicate α_t^H and α_t^L, respectively. The band shows the 95% credible interval approximated with two times posterior standard deviations. Graph (b) presents an average measure of returns-to-scale, calculated as $\alpha^H \times \bar{k}_{pc} + \alpha^L \times (1 - \bar{k}_{pc})$, where $\bar{k}_{pc}$ denotes the percentile position of the estimated threshold over time. For each year t, we compute the relative rank of the convexity-concavity threshold as $\bar{k}_{pc} = \hat{F}_k(\hat{\bar{k}}_t)$, where $\hat{F}_k$ is the empirical cumulative distribution function of total capital, and $\bar{k}_t$ is the cross-sectional convexity-concavity threshold using Bayesian Markov Chain Monte Carlo (MCMC) changepoint estimation method. Graph (c) provides a measure of the relative importance of convexity versus concavity in production, calculated as the ratio $\frac{\text{convexity}}{\text{concavity}} = \frac{\alpha^H \times \bar{k}_{pc}}{\alpha^L \times (1-\bar{k}_{pc})}$. The model specification is shown in Equation (11) with y being total output (Compustat data item *SALE*) and k the sum of physical capital (Compustat data item *PPENT*) and intangible capital measured by Peters and Taylor (2017). We correct for the possible endogeneity issue by using Olley and Pakes (1996) approach. Data is obtained from *Compustat*.

(a) degree of return-to-scale



(b) average return-to-scale



(c) convexity/concavity

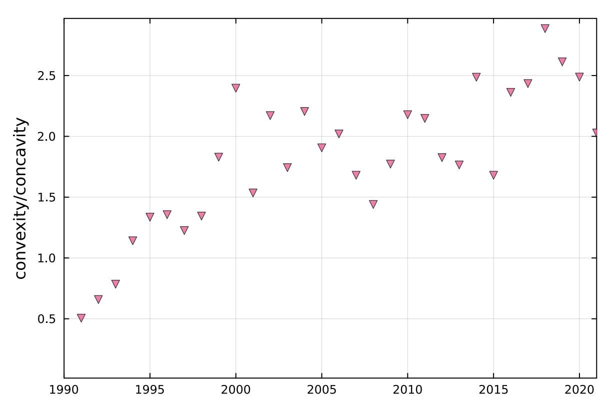


Figure A8: **Long-run Changes in Corporate Production Function with Yu and Phillips (2018) approach**

Notes: Graph (a) shows the estimated cross-sectional convexity-concavity threshold using Yu and Phillips (2018) estimation method. The blue solid line indicates the time-series in natural log scale, while the red dashed line shows the same values in million U.S. dollars. Graph (b) shows the estimated degree of return-to-scale, and the orange and green lines indicate α_t^H and α_t^L, respectively. Data is obtained from *Compustat*.

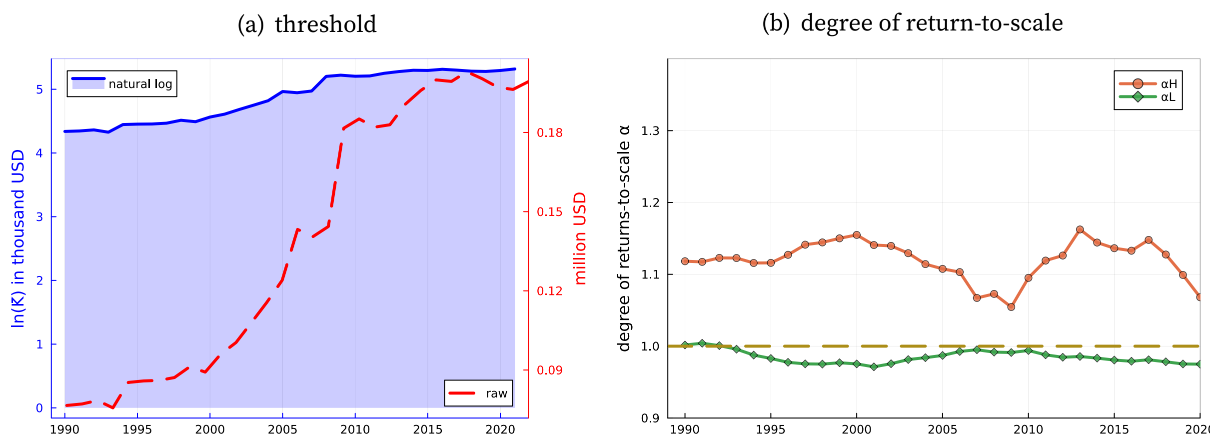
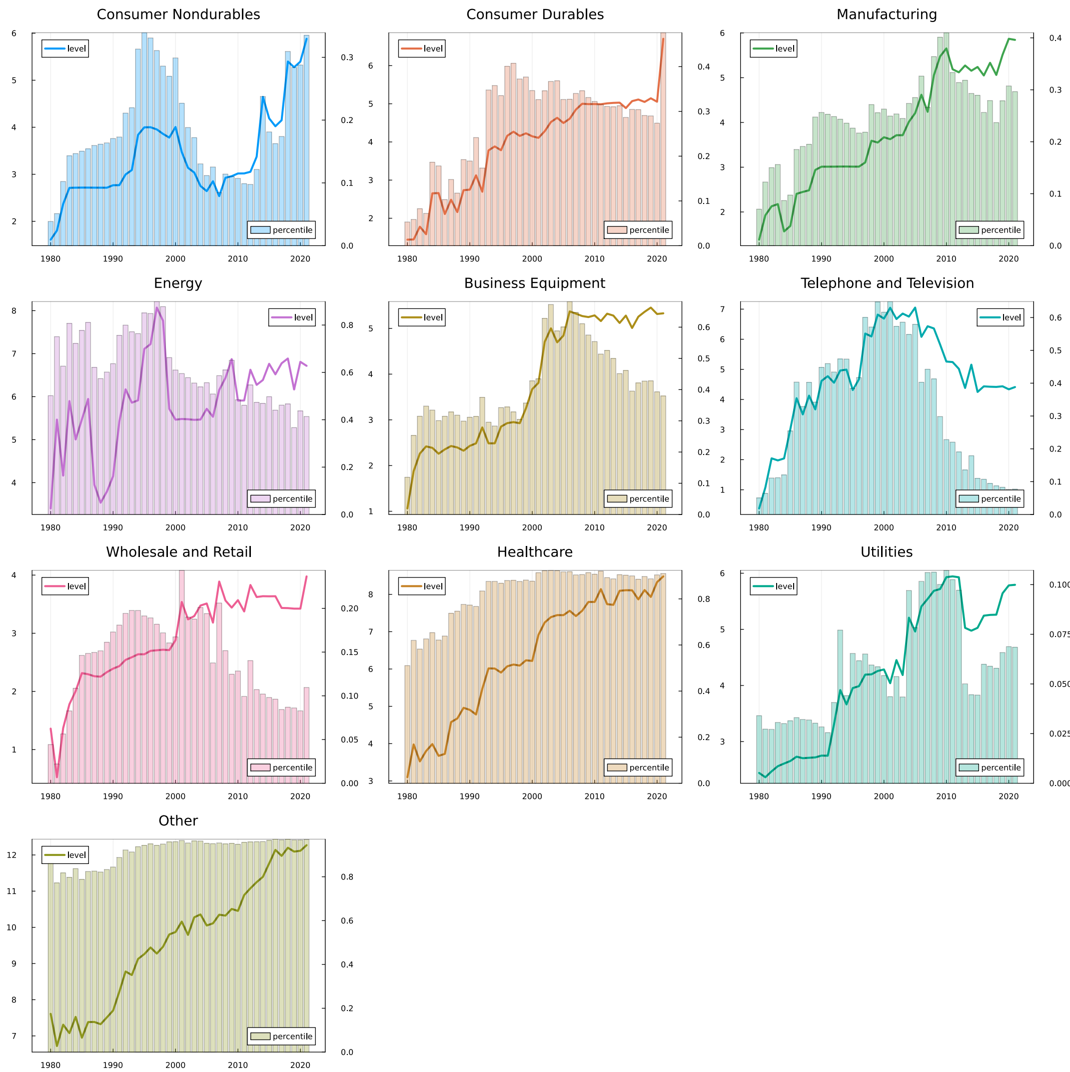


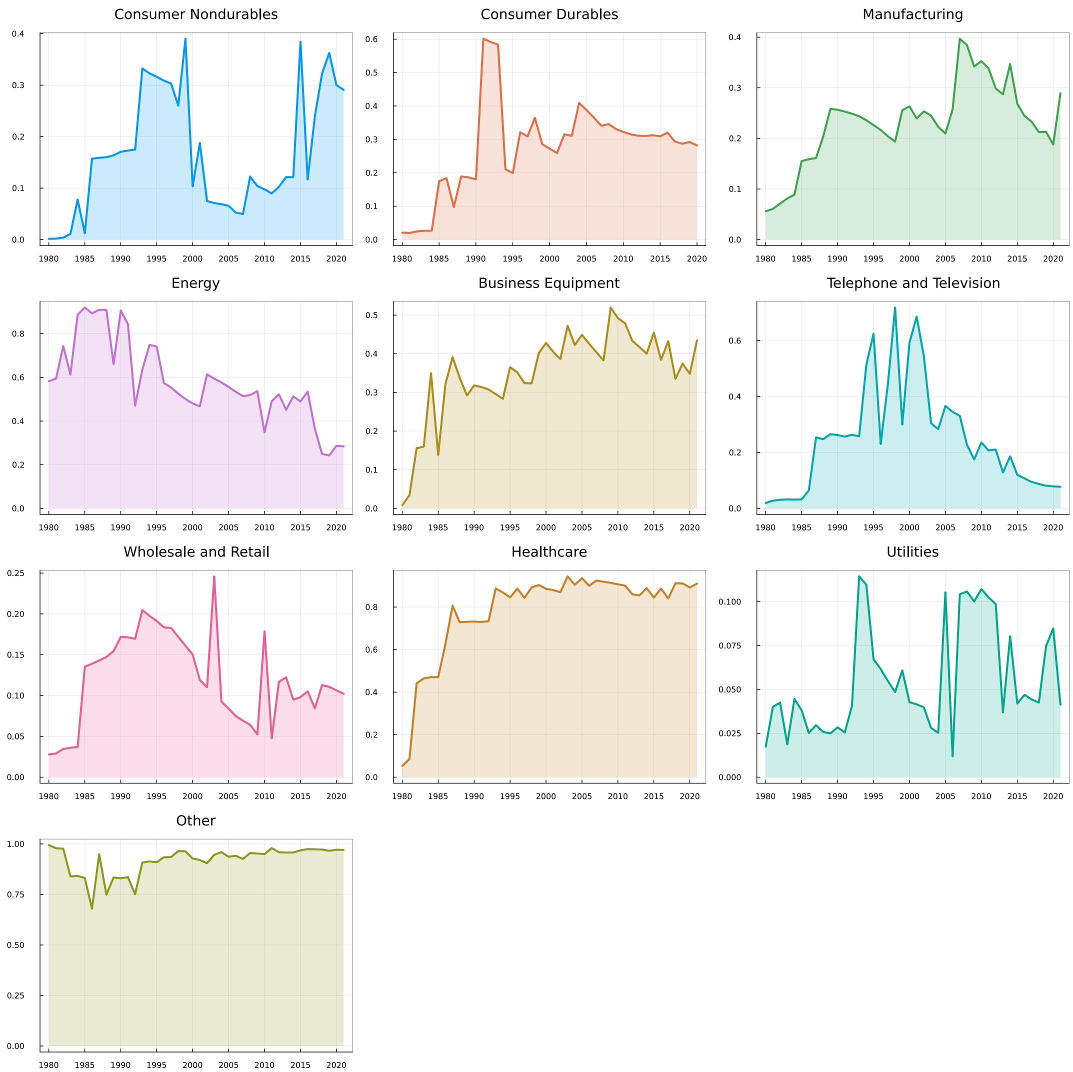
Figure A9: **Long-run Changes in Corporate Production Function: Industry-level Evidence**

Notes: Panel (a) shows the estimated cross-sectional convexity-concavity threshold using Bayesian Markov Chain Monte Carlo (MCMC) changepoint estimation method for different industries. Panel (b) presents the percentile position of the estimated threshold over time. For each year t, we compute the relative rank of the convexity-concavity threshold as $\bar{k}_{pc} = \hat{F}_k(\hat{\bar{k}}_t)$, where $\hat{F}_k$ is the empirical cumulative distribution function of total capital. Panel (c) provides a measure of the relative importance of convexity versus concavity in production, calculated as the ratio $\frac{\text{convexity}}{\text{concavity}} = \frac{\alpha^H \times \bar{k}_{pc}}{\alpha^L \times (1-\bar{k}_{pc})}$. Total output y is measured as Compustat data item *SALE* and capital stock k is the sum of physical capital (Compustat data item *PPENT*) and intangible capital, where we measure the stock of intangible capital by following Peters and Taylor (2017). The industry classification follows Fama-French 10 industry definitions. Data is obtained from *Compustat*.

(a) threshold level



(b) threshold in percentile



(c) convexity/concavity

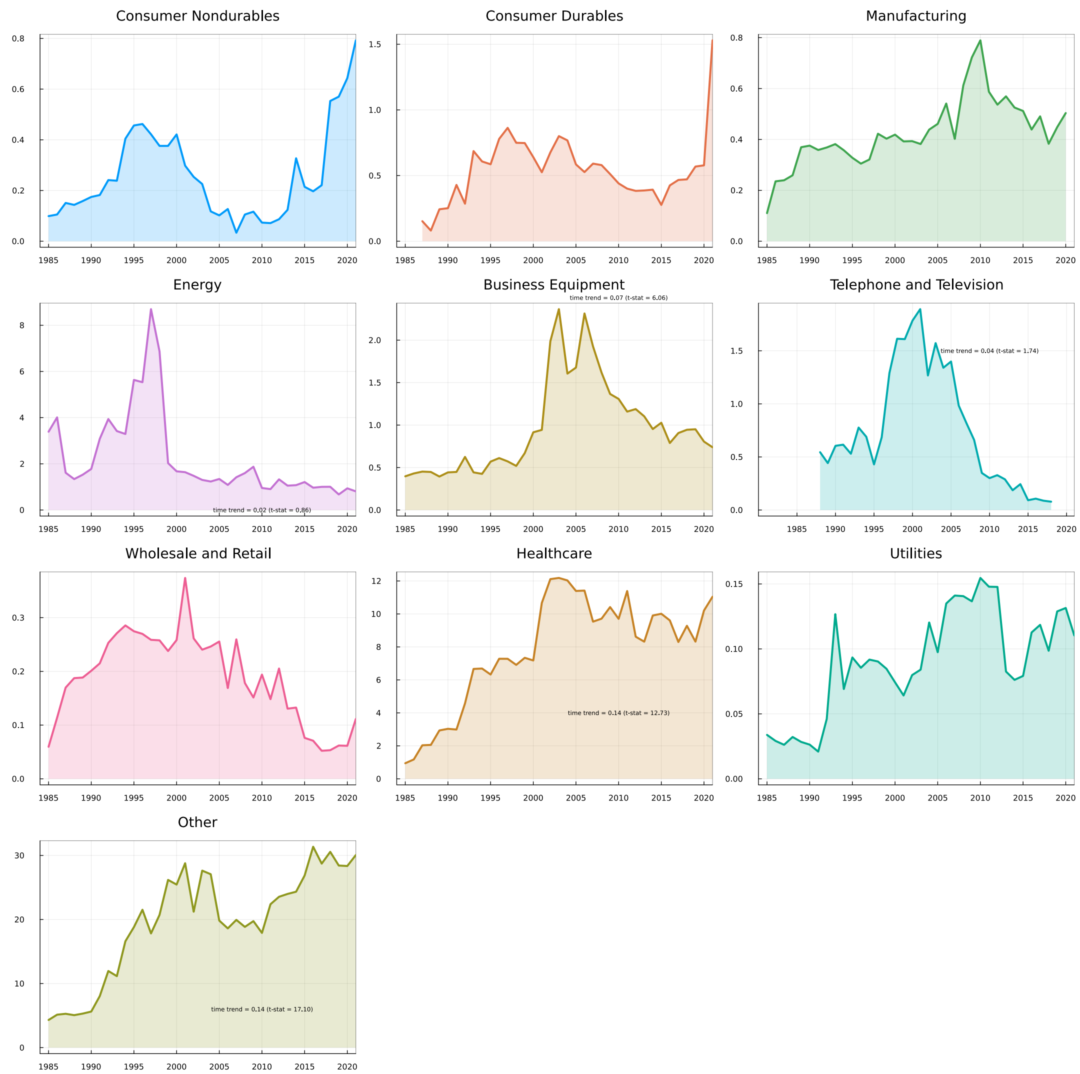


Figure A10: **Correlation between markup and netearnings**

Notes: This figure presents the time-series plot of the cross-sectional correlation between markup and net earnings or intangible capital investment. The shaded area represents the 99% confidence intervals. Data is obtained from *Compustat* and definitions of markup and intangible capital investment are the same as before.

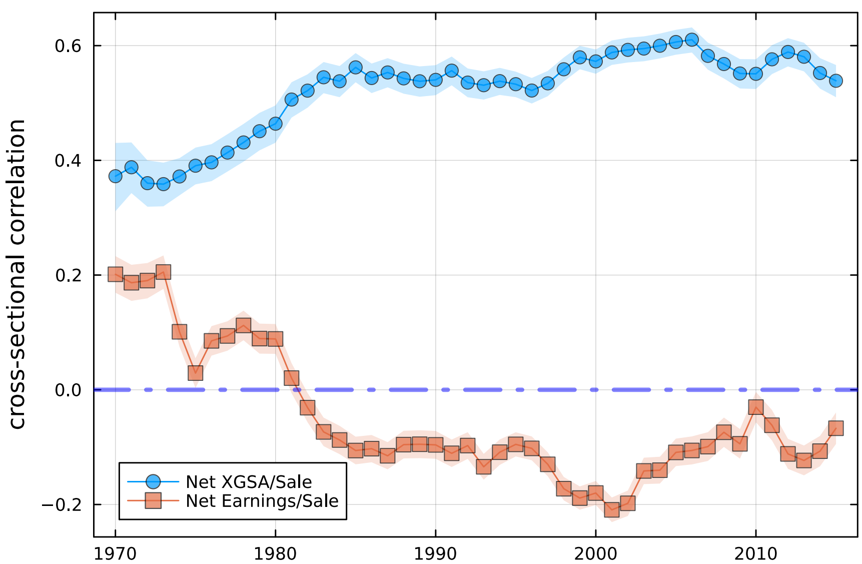


Table A2: **Top 50 Companies with Negative Net Earnings in 2019**

Company Name	Net Earnings (in million US dollars)	Market Capitalization (in million US dollars)	Industry
Boeing Co	-636	183373.2	Manufacturing
Vanjia Corp	-0.041	122949	Construction
General Electric Co	-4979	97520.92	Public Administration
Altria Group Inc	-1293	92731.88	Manufacturing
Tesla Inc	-862	75717.73	Manufacturing
Uber Technologies Inc	-8506	51054.09	Transportation and Public Utilities
Dun & Bradstreet Corp	-560	45586.05	Services
Workday Inc	-480.674	42780.25	Services
Dow Inc	-1359	40582.24	Manufacturing
Occidental Petroleum Corp	-667	36846.36	Mining
Constellation Brands Inc	-11.8	32946.64	Manufacturing
MercadoLibre Inc	-171.999	28431.14	Services
Splunk Inc	-336.668	24498.16	Services
Snap Inc	-1033.66	23119.95	Services
Weyerhaeuser Co	-76	22508.06	Manufacturing
Corteva Inc	-959	22127.94	Agriculture, Forestry and Fishing
Palo Alto Networks Inc	-81.9	21929.07	Services
Halliburton Co	-1131	21484.66	Mining
Hess Corp	-408	20374.04	Mining
Seagen Inc	-158.65	19652.04	Manufacturing
Freeport-McMoRan Inc	-239	19037.12	Mining
Concho Resources Inc	-705	17311.63	Mining
Equifax Inc.	-398.8	16982.54	Services
Roku Inc	-59.937	16054.21	Manufacturing
OKTA INC	-208.913	15703.8	Services
Live Nation Entertainment Inc	-4.882	15273.85	Services
Biomarin Pharmaceutical Inc	-23.848	15205.3	Manufacturing
RingCentral Inc	-53.607	14664.17	Services
Lumen Technologies Inc	-5269	14399.67	Transportation and Public Utilities
DocuSign Inc.	-208.359	14230.25	Services
Western Digital Corp	-754	14027.25	Manufacturing
Exact Sciences Corporation	-83.993	13652.45	Services
Twilio Inc	-307.063	13603.23	Services
Hologic Inc	-203.6	13515.42	Manufacturing
Annaly Capital Management Inc	-2162.865	13471.6	Finance, Insurance and Real Estate
Icahn Enterprises LP	-1098	13165.86	Public Administration
Lyft Inc	-2602.241	13017.68	Transportation and Public Utilities
CrowdStrike Holdings Inc	-141.779	13008.99	Services
Alnylam Pharmaceuticals Inc	-886.116	12920.69	Manufacturing
Noble Energy Inc	-1512	12045.76	Mining
Slack Technologies Inc	-571.058	11512.61	Services
Equitable Holdings Inc	-1733	11490.76	Finance, Insurance and Real Estate
Datadog Inc	-16.71	11197.5	Services
Zscaler Inc	-28.655	10723.61	Services
Formula One Group - The Liberty Media Group	-311	10647.7	Services
Chewy Inc	-252.37	10640.27	Retail Trade
Pinterest Inc	-1361.371	10623.01	Services
Coupa Software Inc	-90.832	10398.85	Services
Coty Inc	-3784.2	10106.28	Manufacturing
Darden Restaurants Inc	-52.4	9983.653	Retail Trade