Markups, Production Technology, and the Cost of Capital in Financial Markets

Ginha Kim
James Traina*

June 2025

Abstract

We develop a new approach to measuring markups by combining cost share and implied cost of capital techniques. This method resolves identification problems in production function estimation. Using data on US public firms from 1970 to 2022, we find average markups remain stable and close to competitive levels. Heterogeneity-adjusted markups are lower than conventional estimates. Markup dispersion increases substantially, with the 99th percentile rising from 1.10 to 1.25. We document a strong negative relationship between markups and cost of equity capital. These patterns support superstar hypotheses and suggest that financing advantages and technological change drive increasing heterogeneity across firms.

JEL Codes E22, E44, D24, G31, G32

Keywords Cost of Capital; Investment; Markups; Production Function Estimation; Firm Heterogeneity

*Kim: Chicago Booth, ginhaki1@chicagobooth.edu. Traina: NYU Stern Abu Dhabi, james.traina@nyu.edu. We thank Arshia Hashemi and Uyen Tran for their valuable help and feedback.

Production-based markup estimates following (Hall, 1988) largely overlook differences in financing costs when calculating the user cost of capital, though finance research documents large heterogeneity in risk premia and required returns (Fama and French, 1992; Cochrane, 2009). Motivated by this gap in the literature, we ask: How does incorporating firm-specific financing costs change our understanding of markup trends and firm heterogeneity? Firms pay different costs to finance their capital. A stable manufacturer might pay 5% while a growing tech venture might pay 25% because of different risk profiles. These values have also changed markedly from the 1980s to the present (Duarte and Rosa, 2015). This variation matters because underestimating costs leads to overestimating markups, directly affecting our measurement of market power.

We address our research question by synthesizing insights from two literatures—one in productivity and macroeconomics, the other in accounting and finance—to measure user costs and markups in Compustat microdata. We work from the cost share approach to production function estimation, which equates elasticities to input cost shares under general conditions and avoids the severe non-identification problems common in other approaches (Gandhi et al., 2020; Bond et al., 2021). A key challenge is measuring the user cost of capital, particularly the risk premia embedded in the cost of equity capital. To overcome this challenge, we recover the implied cost of equity capital by applying the present value relation to market values and expected future profits from an earnings forecast model, following Gebhardt et al. (2001) and Hou et al. (2012).[1] Our contribution is to combine these two flexible and empirically successful methodologies, allowing rich firm-year-level heterogeneity in costs of capital, production technologies, and markups for nearly all US public firms.

Our synthesis alters the narrative of broadly rising market power. When accounting for firm-specific capital costs, average markups for US public firms remain close to 1 throughout our sample period (1970 to 2022), with mean and median markups stable between 1.00 and 1.05, indicating broadly competitive pricing. However, we document a significant increase in the dispersion of markups across firms consistent with the superstar hypothesis of Autor et al. (2020)—where economic changes create winner-take-all effects that favor the most productive firms. Supporting this view, we find a strong negative relationship between markups and the implied cost of equity capital, suggesting an advantage in financial funding might be a key mechanism that generates superstars (Liu et al., 2022). Our findings also suggest technological change, not declining competition, drives the increasing dispersion in markups, which may cause welfare losses through misallocation (Baqaee and Farhi, 2020; Edmond et al., 2023).

We develop a unified model linking investment decisions, capital costs, and markups. Building on the dynamic cost minimization framework of Basu and Fernald (2001), we extend it to incorporate how investors value expected future profits and connect firm behavior to investor-required returns in financial markets. Consider the thought experiment of holding fixed a firm's choice of labor

[1] Pástor et al. (2008) and Frank and Shen (2016) show this method successfully predicts the positive risk-return relationship and negative investment-cost relationship, respectively. See Lee et al. (2021) for a review in the finance and accounting literatures.

and investment. Mechanically, a higher cost of equity capital implies a higher user cost of capital (Jorgenson, 1963), which implies a higher capital cost share, a higher estimated output elasticity of capital, a lower estimated output elasticity of labor, and a lower estimated markup. Firms also respond to a higher user cost by decreasing their investment rates and capital stocks, so the resulting bias depends on endogenous responses.

We then introduce our finance-based approach to estimating firm-specific user costs of capital, production technologies, and markups. We apply the implied cost of equity capital method developed by Gebhardt et al. (2001) and extended by Hou et al. (2012) to estimate risk-adjusted discount rates for individual firms. We adjust for differences in leverage, depreciation, and taxes in the user cost formula. Our method only requires firms to have publicly traded equity and data on basic accounting items, allowing us to construct a panel of over 15,000 public firms from 1970 through 2022. These estimates display significant variation across firms and over time. We then combine these capital-cost estimates with the cost-share approach to production function estimation to recover firm-year-level production technologies and markups.

Our estimates reveal several key patterns. First, large variation exists in the implied cost of equity capital—both across time and in the cross-section of firms. The interquartile difference of the implied cost of equity capital has varied from 2.5 to 6.3 percentage points over time. Second, high implied costs of equity capital strongly predict low markups—a percentage point increase in the implied cost of capital is associated with 0.5% lower markups. Firms don't fully offset higher costs of capital with proportionally lower capital stocks. Third, once we incorporate firm-specific capital costs, average markups remain stable around 1.00 from the 1980s onward, showing no significant aggregate increase.[2] We find that ICC-adjusted markups are approximately 15% lower than conventional estimates that use a common risk-free rate for all firms, with this bias growing over time as financing cost heterogeneity has increased.

Fourth, we estimate a sizable rise in markup dispersion; average markups of the top quartile increased from 1.03 to 1.14 over 1980 to 2022, while the bottom quartile has declined from 0.90 to 0.80 over the same period. The markup at the 99th percentile shows an especially large rise from 1.10 in 1980 to 1.25 in 2022. On average, these firms that charge markups above these levels face a 6.5% cost of equity capital, which is considerably lower relative to the overall median of 9% and mean of 10%. Finally, production-based markup estimates prove sensitive to assumptions about capital costs. Using fixed values for implied cost of capital across industry or years, we find that ignoring firm-year level heterogeneity in user costs leads to lower markup estimates.

Our first contribution is methodological: we develop a new approach to measuring markups that combines the cost shares method with firm-specific financing cost estimates from the finance and accounting literature. Cost share estimators are quickly becoming a common workhorse model

[2] Our method assumes cost minimization rather than profit maximization, which allows for overproduction relative to the profit-maximizing level and therefore markups below 1. This can occur due to empire-building managerial incentives, agency problems, or other organizational frictions.

for markup estimation, as they avoid the severe nonidentification problems of Olley and Pakes (1996)-style estimators documented by Bond and Söderbom (2005), Gandhi et al. (2020), Klette and Griliches (1996), and Bond et al. (2021). Our finance-based approach solves one of the thorniest problems in applying cost shares—measuring the user cost of capital. Two key advantages distinguish our approach. First, we estimate the user cost of capital individually for each firm, capturing granular variation in financing costs that standard approaches miss. Second, because we build on the cost shares framework, our method flexibly accommodates heterogeneity in production technologies and markups across firms. This flexibility matters for understanding the cross-sectional distribution of markups and firm heterogeneity, which is a central theme in the literature on market power and firm dynamics (Syverson, 2011; Autor et al., 2020; Van Reenen, 2018). By solving these measurement issues, we directly address concerns raised by Basu (2019) and Syverson (2019) about potential biases in markup estimates.

Our second contribution is empirical: we document patterns in firm heterogeneity that challenge existing narratives about rising market power. When accounting for firm-specific capital costs, we find average markups for US public firms remain close to 1 throughout 1970 to 2022, with mean and median values stable between 1.00 and 1.05, indicating that markets have remained broadly competitive. Taken together with results on concentration and profits (Grullon et al., 2019; Barkai, 2020; Covarrubias et al., 2020; Davis et al., 2024), we show the markups hypothesis has conflicting evidentiary support. We also show that properly accounting for user costs matters greatly for markup estimates (Karabarbounis and Neiman, 2019; De Loecker et al., 2020; Farhi and Gourio, 2018).

While averages remain stable, we observe a significant increase in markup dispersion across firms, with top-quartile firms showing rising markups while bottom-quartile firms show declining markups. This pattern is consistent with the superstar hypothesis proposed by Autor et al. (2020) and Van Reenen (2018) and highlights microdata heterogeneity that was a central theme in Syverson (2011). Markup dispersion also signals misallocation, implying large welfare losses (Baqaee and Farhi, 2020; Edmond et al., 2023). We also document a strong negative relationship between markups and the implied cost of equity capital, suggesting that financing advantages might be the mechanism that generates these superstars (Liu et al., 2022). Because only systematic risks that investors cannot diversify away enter the user cost, markups may still compensate idiosyncratic risks borne by entrepreneurs (Boar et al., 2022; Di Tella et al., 2024). Our empirical evidence is also consistent with findings from David and Venkateswaran (2019) and Foster et al. (2022) where technological heterogeneity drives increasing dispersion in capital allocation and firm performance.

Our third contribution examines the role of intangible capital in markup patterns, contributing to the literature on how intangibles affect market structure and firm dynamics (Crouzet and Eberly, 2023). We find no significant relationship between intangible asset intensity and markups. R&D intensity shows a strong negative relationship with markups, suggesting that innovation-intensive firms face higher costs that compress margins. Our findings indicate that while intangibles may contribute to firm heterogeneity, they do not mechanically translate into higher markups, showing that properly measuring capital costs matters when studying intangible-intensive firms.

The remainder of the paper is organized as follows: Section 1 develops our model that derives the relationship between markups, production technology, and the cost of capital in financial markets. Section 2 details our empirical implementation on the Compustat sample. Section 3 presents our empirical results, and Section 4 concludes.

1 A Unified Model of Investment and Markups

This section develops a framework that connects investment decisions, markups, and capital costs. Our starting point is the investment literature tracing back to Jorgenson (1963) and Hayashi (1982). We draw on insights from Cooper and Ejarque (2001), Abel and Eberly (2011), and Balvers et al. (2017) to highlight how markups or concavity in the profit function can affect investment incentives, even with frictionless capital processes. Our approach also builds on the production-based asset pricing literature (Cochrane, 1991) by incorporating financial market equilibrium conditions. We begin by modeling how firms dynamically minimize costs as in Basu and Fernald (2001). We then introduce market power via markups and show how these markups relate to marginal cost, output elasticities, and revenue shares (Hall, 1988). Under constant returns to scale, labor and capital cost shares reveal their underlying production elasticities. Next, we derive the user cost of capital from the firm's envelope condition and investment equations, linking it to depreciation and risk. Finally, we connect these results to asset prices by comparing investment returns, financial-market returns, and their risk premia in capital-market equilibrium.

1.1 A Dynamic Cost-Minimization Problem

We consider a panel of firms $i = 1, 2, \dots$ operating in discrete time $t = 0, 1, 2, \dots$. Each firm i produces output Y_{it} using capital K_{it} and labor L_{it} according to a production function:

$$Y_{it} = F(K_{it}, L_{it}, t) \tag{1.1}$$

where the time argument captures possible technological change and other time-varying factors. The production function exhibits constant returns to scale, consistent with evidence from industry-level (Basu and Fernald, 1997) and establishment-level (Syverson, 2004) data. Capital evolves through investment according to:

$$K_{it+1} = (1 - \delta_{it})K_{it} + I_{it} \tag{1.2}$$

where $\delta_{it} \in (0, 1)$ is the firm and time-specific depreciation rate and I_{it} is gross investment.

Each firm minimizes the present value of costs—labor plus investment—across an infinite horizon, discounting future outlays by a stochastic discount factor M_{t+1}. Such an approach follows the style of Cochrane (1991) and the subsequent production-based asset pricing literature. The firm's value function $V(K_{it}, t)$ represents the minimum expected discounted cost:

$$V(K_{it}, t) = \min_{L_{it}, I_{it}} \Big\{ W_{it}L_{it} + I_{it} + \mathbb{E}_t\big[M_{t+1}V(K_{it+1}, t+1)\big] \Big\} \tag{1.3}$$

subject to $Y_{it} = F(K_{it}, L_{it}, t)$ and the capital law of motion in (1.2). Here, W_{it} is the wage rate, and $\mathbb{E}_t[\cdot]$ denotes expectations conditional on information at time t. As additional capital reduces future production costs, we have $\partial V/\partial K_{it} < 0$.

1.2 First-Order Conditions

To find optimal input choices, we form the Lagrangian with multiplier λ_{it} on the production constraint:

$$\mathcal{L}_{it} = W_{it}L_{it} + I_{it} + \mathbb{E}_t\Big[M_{t+1}V(K_{it+1}, t+1)\Big] + \lambda_{it}\Big[Y_{it} - F(K_{it}, L_{it}, t)\Big] \tag{1.4}$$

The multiplier λ_{it} represents the marginal cost of production. The firm's optimization yields two conditions that characterize optimal behavior.

Differentiating the Lagrangian with respect to L_{it}:

$$\frac{\partial \mathcal{L}_{it}}{\partial L_{it}} = W_{it} - \lambda_{it}\frac{\partial F}{\partial L} = 0$$

Rearranging:

$$W_{it} = \lambda_{it}\frac{\partial F}{\partial L} \tag{2.1}$$

The firm hires labor until the wage equals the value of labor's marginal product.

Differentiating with respect to I_{it} and considering $\frac{\partial K_{it+1}}{\partial I_{it}} = 1$:

$$\frac{\partial \mathcal{L}_{it}}{\partial I_{it}} = 1 + \mathbb{E}_t\Big[M_{t+1}\frac{\partial V}{\partial K_{it+1}}\Big] = 0$$

This simplifies to:

$$1 = -\,\mathbb{E}_t\Big[M_{t+1}\frac{\partial V}{\partial K_{it+1}}\Big] \tag{2.2}$$

The firm invests until the expected discounted marginal benefit of additional capital equals its unit cost.

To derive how the value function changes with current capital, we apply the envelope theorem:

$$\frac{\partial V}{\partial K_{it}} = \frac{\partial \mathcal{L}_{it}}{\partial K_{it}} = -\,\lambda_{it}\frac{\partial F}{\partial K} + (1-\delta_{it})\,\mathbb{E}_t\Big[M_{t+1}\frac{\partial V}{\partial K_{it+1}}\Big] \tag{2.3}$$

We assume depreciation is known at time t, hence outside the expectation. This equation captures both the immediate production benefit of capital and its future value after depreciation.

Substituting the investment FOC (2.2):

$$\frac{\partial V}{\partial K_{it}} = -\lambda_{it}\frac{\partial F}{\partial K} - (1-\delta_{it}) \tag{2.4}$$

This connects the current shadow value of capital to its productive contribution and continuation value. Intuitively, the marginal value of capital depends negatively on the marginal product (since more capital lowers marginal cost) and positively on its continuation value after depreciation.

1.3 Markups

Cost minimization is broadly compatible with many market structures including various forms of imperfect competition. Firms minimize cost given whatever output price they face, even under monopoly power.

The markup is the ratio of output price to marginal cost:

$$\mu_{it} = \frac{P_{it}}{\lambda_{it}} \tag{3.1}$$

From the labor FOC (2.1), we have $\lambda_{it} = \frac{W_{it}}{\partial F/\partial L}$. Substituting:

$$\mu_{it} = \frac{P_{it}\,\frac{\partial F}{\partial L}}{W_{it}} \tag{3.2}$$

Multiplying both the numerator and denominator by L_{it}/Y_{it}:

$$\mu_{it} = \left(\frac{\partial \ln F}{\partial \ln L}\right)\left(\frac{P_{it}Y_{it}}{W_{it}L_{it}}\right) \tag{3.3}$$

The markup equals the ratio of labor's output elasticity to labor's revenue share. When $\mu_{it} = 1$, the firm behaves competitively. Values above 1 indicate market power, where firms price above marginal cost because they face downward-sloping (residual) demand curves. Markups can also fall below 1. Real-world factors (e.g., regulatory constraints, strategic underpricing, long-term contracts, empire building) might push firms to price below statically profit-maximizing levels.

This ratio approach to measuring markups follows Hall (1988), who first showed how researchers can infer markups from production data using cost-minimization conditions. Recent empirical work by De Loecker et al. (2020) has documented significant increases in market power applying this insight to microdata, though this finding remains debated (Traina, 2018; Syverson, 2019; Basu, 2019).

1.4 Cost Shares

Under constant returns to scale, a fundamental relationship emerges between factor cost shares and output elasticities. Euler's theorem gives:

$$F(K_{it}, L_{it}, t) = K_{it}\,\frac{\partial F}{\partial K} + L_{it}\,\frac{\partial F}{\partial L} \tag{4.1}$$

We define the total economic cost in period t as:

$$\mathcal{C}_{it} = W_{it}L_{it} + R_{it}K_{it} \tag{4.2}$$

where R_{it} is the user cost of capital. This formulation of the user cost concept builds on Jorgenson (1963), who showed how interest rates, depreciation, and (in later work) taxes determine capital's implicit rental price. The user cost of capital measures the total cost incurred by using an additional unit of capital. We introduce it explicitly to quantify capital costs consistently across periods.

From cost minimization:

$$W_{it} = \lambda_{it}\frac{\partial F}{\partial L}, \quad R_{it} = \lambda_{it}\frac{\partial F}{\partial K} \tag{4.3}$$

Multiplying Euler's theorem (4.1) by λ_{it} and substituting (4.3):

$$\lambda_{it}\,F(K_{it}, L_{it}, t) = \lambda_{it}K_{it}\,\frac{\partial F}{\partial K} + \lambda_{it}L_{it}\,\frac{\partial F}{\partial L} = R_{it}K_{it} + W_{it}L_{it} = \mathcal{C}_{it} \tag{4.4}$$

Since $F(K_{it}, L_{it}, t) = Y_{it}$, we have $\lambda_{it}Y_{it} = \mathcal{C}_{it}$.

Dividing each factor's cost by total cost gives:

$$\frac{W_{it}L_{it}}{\mathcal{C}_{it}} = \frac{\lambda_{it}L_{it}}{\lambda_{it}Y_{it}}\frac{\partial F}{\partial L} = \frac{L_{it}}{Y_{it}}\frac{\partial F}{\partial L} = \frac{\partial \ln F}{\partial \ln L} \tag{4.5}$$

$$\frac{R_{it}K_{it}}{\mathcal{C}_{it}} = \frac{\lambda_{it}K_{it}}{\lambda_{it}Y_{it}}\frac{\partial F}{\partial K} = \frac{K_{it}}{Y_{it}}\frac{\partial F}{\partial K} = \frac{\partial \ln F}{\partial \ln K} \tag{4.6}$$

This equivalence between factor shares and output elasticities holds regardless of market structure. Under constant returns to scale, each factor's share in total cost equals its output elasticity, regardless of markup values.

1.5 User Cost of Capital

We now derive the user cost of capital from our dynamic framework. Starting with the envelope condition (2.3):

$$\frac{\partial V}{\partial K_{it}} = -\,\lambda_{it}\frac{\partial F}{\partial K} + (1-\delta_{it})\,\mathbb{E}_t\Big[M_{t+1}\frac{\partial V}{\partial K_{it+1}}\Big]$$

Substituting the investment FOC (2.2):

$$\frac{\partial V}{\partial K_{it}} = -\,\lambda_{it}\frac{\partial F}{\partial K} - (1-\delta_{it}) \tag{5.1}$$

Rearranging and using $R_{it} = \lambda_{it}\,\frac{\partial F}{\partial K}$ from (4.3):

$$R_{it} = -\,\frac{\partial V}{\partial K_{it}} - (1-\delta_{it}) \tag{5.2}$$

This formulation shows how the user cost relates to the firm's shadow value of capital (capturing the savings on future costs due to additional capital), adjusted for depreciation.

From the envelope condition at time $t+1$:

$$\frac{\partial V}{\partial K_{it+1}} = -\,\lambda_{it+1}\frac{\partial F}{\partial K} - (1-\delta_{it+1})$$

Substituting into the investment FOC and using $R_{it+1} = \lambda_{it+1}\frac{\partial F}{\partial K}$:

$$1 = \mathbb{E}_t\Big[M_{t+1}\big(R_{it+1} + (1-\delta_{it+1})\big)\Big] \tag{5.3}$$

Defining the investment return as $R^I_{it+1} \equiv R_{it+1} + (1-\delta_{it+1})$:

$$\mathbb{E}_t\big[M_{t+1}\,R^I_{it+1}\big] = 1 \tag{5.4}$$

This equation is a standard asset-pricing Euler equation stating that the discounted expected return on investing in capital must equal its marginal cost now. Investors will invest until their expected discounted marginal benefit equals their marginal cost, i.e., until there are no arbitrage opportunities in equilibrium.

1.6 Market Values and Expected Returns

Market outcomes and firm decisions align through pricing relationships that connect capital markets to firm optimization. The investment return derived in (5.4):

$$R^I_{it+1} = R_{it+1} + (1-\delta_{it+1}) \tag{6.1}$$

combines the productive service value R_{it+1} and the undepreciated capital value $(1-\delta_{it+1})$. The investment FOC implies:

$$\mathbb{E}_t\big[M_{t+1}\,R^I_{it+1}\big] = 1 \tag{6.2}$$

Similarly, the firm's equity return R^E_{it+1} must satisfy:

$$\mathbb{E}_t\big[M_{t+1}\,R^E_{it+1}\big] = 1 \tag{6.3}$$

Both equations share the same form, reflecting how capital market equilibrium eliminates persistent differences between investment and equity returns. Put plainly, if equity had systematically higher returns than physical investment (or vice versa), investors would shift resources accordingly until returns equalize. Firms themselves can equivalently invest internally or externally, eliminating persistent differences.

For any return, we can separate its expected value into risk-free and risk components:

$$\mathbb{E}_t[R^E_{it+1}] = R^F_t + \theta_{it} \tag{6.4}$$

where θ_{it} is the risk premium determined by the covariance of returns with the stochastic discount factor. Assets positively correlated with the stochastic discount factor offer insurance and command lower returns; negatively correlated assets, bearing systematic risk, require higher expected returns.

Both investment and equity returns derive from the same underlying assets and cash flows, so their expected values converge in equilibrium. This convergence results from no-arbitrage conditions requiring that returns on assets yielding identical future payoffs must align, consistent with standard asset-pricing theory.

Defining the required return $\rho_{it} = \mathbb{E}_t[R^I_{it+1}] - 1$, we get:

$$R_{it} = \rho_{it} + \delta_{it} \tag{6.5}$$

This formula holds across market structures because capital acquisition occurs in competitive markets separate from product markets. Economic rents from market power accrue to truly scarce factors—patents, brand assets, or entry barriers—rather than to competitively supplied capital inputs.

This theoretical framework guides our empirical approach. The markup equation (3.3) connects to observable firm behavior—for example, a technology firm with high fixed costs but low marginal production costs (like software companies) would show a high ratio of revenues to labor costs, indicating large markup power. The user cost formula (5.2) provides the foundation for our measurement using financial statement data, where we will estimate depreciation rates and risk premia from observable company accounts and market prices. The link between capital market returns and investment returns established in equations (6.2) and (6.3) explains why we can use market-based measures of expected returns in our production-based markup estimates.

2 A Finance Approach to Markup Estimation

Having established our framework linking investment decisions, capital costs, and markups, we now implement this approach using financial statement data. We examine over five decades of standardized accounting information for US public companies, constructing a panel that allows us to track markup patterns across firms, industries, and time. De Loecker et al. (2020) document the rise of market power using similar production-based techniques, but our approach differs by incorporating financial market information to estimate capital costs.

Our empirical method follows a sequential process: we first develop a cross-sectional statistical model to forecast firm-level earnings, then use these projections to estimate the cost of equity, and finally calculate markups based on the resulting user cost, implied production technology, and observed sales and expenditures. Our approach allows us to measure markups without requiring detailed product-level information, instead using readily available financial statements.

2.1 Terminology and Notation

When analyzing capital costs, different fields often use the same term—"cost of capital"—to mean different things. To avoid confusion, we use specific terms for each component:

- **User Cost of Capital** (R_{it}): The total economic cost of using capital, which includes financing, depreciation, and tax effects
- **Weighted Average Cost of Capital** (WACC or R_{it}^{A}): The cost of financing through both debt and equity
- **Cost of Equity Capital** (R_{it}^{E}): The required return to equity investors (also called the implied cost of capital or ICC in some literature)
- **Cost of Debt Capital** (R_{it}^{D}): The interest rate paid on borrowed funds

These components relate through two simple equations:

$$\text{User Cost} = \text{WACC} + \text{Depreciation Rate} + \text{Tax Adjustment} \tag{1}$$

$$\text{WACC} = (\text{Equity Share} \times \text{Cost of Equity}) + (\text{Debt Share} \times \text{Cost of Debt}) \tag{2}$$

Our theoretical model showed a simplified version where user cost equals expected return plus depreciation. Our empirical implementation extends this to include financing structure and taxes.

2.2 Data

We obtain accounting data for US public firms from Compustat North America starting in 1960. We report results from 1970 to 2022. Our earnings forecast model requires a 10-year estimation window, so the first forecast uses data from 1960-1969 to predict earnings for 1970. This dataset provides standardized financial statement information for companies trading on major US exchanges. Following standard practice in finance and accounting, we focus on industrial and service firms, excluding financial institutions (SIC 6000-6999) and utilities (SIC 4000-4999) because of their distinct regulatory environments and accounting practices that make direct comparisons with other sectors problematic.

We apply several filters for data quality and consistency. First, we require firms to have non-missing values for variables used to forecast earnings and estimate markups, including total assets, earnings, sales, operating expenses, interest expense, depreciation, and invested capital. Second, we remove observations with zero or negative sales, as these represent non-operational periods or data errors rather than normal business activities. Third, we exclude cases with negative invested capital or operating expenses, which show accounting irregularities. Last, we require positive common equity (CEQ), as negative equity represents financial distress that complicates standard financial analysis.

These filters help our sample comprise firms engaged in standard business operations with reliable financial reporting. We include firms of all sizes and sectors. Overall, Compustat offers a broad view of US corporate activity.

The resulting panel dataset is unbalanced, as firms enter and exit the sample due to new listings, delistings, mergers, acquisitions, and bankruptcies. The number of firms increases markedly during the 1980s and 1990s, reflecting the broader expansion of public equity markets during this period. Our sample peaks in the late 1990s with over 6,000 firms per year, before declining in recent years as merger activity and decreased public listings reduced the number of independent public firms. For example, while we observe nearly 7,000 firms in 1997, this number falls below 3,500 by 2022. This pattern is consistent with documented trends in US equity markets (Kahle and Stulz, 2017).

We define our key variables:

Operating Expenses (XOPR) measure the firm's variable costs of conducting business. As noted by Traina (2018), treatment of this variable significantly affects markup estimates. This category may include research and development (R&D) expenses. We maintain Compustat's classification rather than reclassifying R&D as a capital expenditure to ensure consistent treatment across firms. Our results are robust to alternative treatments that separate SG&A and its R&D subcomponent from operating expenses, as long as we maintain that these are productive inputs.

Invested Capital (ICAPT) represents the total resources allocated to business operations. This measure captures all non-financial assets employed by the firm, including property, plant and equipment, inventories, and recorded intangible capital such as goodwill, patents, and trademarks. Unlike theoretical models that separate tangible and intangible capital, we follow accounting conventions in using this more comprehensive measure of productive assets (Damodaran, 2007; Davis et al., 2024; Ayyagari et al., 2024).

For implied cost of capital estimation, we use Income Before Extraordinary Items (IB) as our earnings measure. Dividends (DV) capture annual shareholder distributions. Total Assets (AT) measure firm size. Accruals (AC) capture the difference between reported income and actual cash flows. Before 1988, we use the balance sheet method, calculating accruals as changes in non-cash current assets minus current liabilities (excluding short-term debt and taxes payable) minus depreciation and amortization. From 1988 on, we use the cash flow statement method, calculating accruals as net income minus cash flow from operations following Hou et al. (2012). We also use

Common Equity (CEQ) for book value of equity and calculate market value of equity as stock price (PRCC) times shares outstanding (CSHO).

For user cost of capital calculations, we additionally require Interest Expense (XINT) to measure debt financing costs, Depreciation (DP) to capture capital consumption, and Income Taxes (TXT) to calculate after-tax costs. These variables, combined with the ICC estimates, allow us to construct firm-specific user costs that account for each firm's unique financing structure, tax situation, and capital depreciation patterns.

Table 1 presents summary statistics for the key variables used in our analysis. We use nominal values as reported, without inflation adjustments, as our analysis focuses primarily on cross-sectional comparisons within years rather than absolute value comparisons across time.

Table 1: Summary Statistics of Compustat Input Variables

	N	Mean	SD	p10	p25	Median	p75	p90
E_t	160,996	107.44	518.05	-12.22	-0.15	3.83	33.60	199.00
A_t	160,996	2,333.52	8,357.49	12.34	38.29	168.43	952.51	4,412.24
D_t	160,996	43.18	218.11	0.00	0.00	0.00	5.38	57.43
DD_t	160,996	0.49	0.50	0.00	0.00	0.00	1.00	1.00
$NegE_t$	160,996	0.26	0.44	0.00	0.00	0.00	1.00	1.00
AC_t	160,996	-115.63	465.57	-220.50	-41.47	-4.50	0.11	6.07
Sales	160,996	2,698.40	14,610.07	12.05	40.89	181.70	930.46	4,130.15
Total variable cost	160,996	2,298.19	12,735.13	11.57	37.41	159.05	796.32	3,480.10
Capital stock	160,996	1,871.12	10,788.32	7.49	24.06	106.28	604.51	2,727.42
Common equity	160,996	1,242.68	7,402.22	5.79	18.56	82.10	417.00	1,745.22
Debt	160,996	894.45	6,740.83	0.00	2.99	24.44	216.76	1,199.94
Interest expense	160,996	43.42	271.07	0.03	0.29	2.00	15.14	75.58
Income taxes	160,996	70.93	567.88	-0.73	0.04	2.27	17.54	93.20
Depreciation	160,996	135.68	829.38	0.37	1.29	6.30	38.42	185.52

Notes: This table presents summary statistics for our sample of non-financial, non-utility US public firms from 1970 to 2022. Values are in nominal terms (millions of dollars). Variables used in the earnings forecast regressions—earnings (E_t), total assets (A_t), dividends (D_t), dividend indicator (DD_t), negative earnings indicator ($NegE_t$), and accruals (AC_t)—are winsorized at the 1st and 99th percentiles by year to mitigate the influence of extreme observations.

In constructing our sample, we make minimal adjustments to the underlying data, preferring to let the empirical patterns emerge naturally. This approach avoids imposing researchers' priors about what constitutes "normal" operations. The long time series also allows us to examine how corporate behavior and financial structures evolve through multiple business cycles, technological changes, and regulatory shifts.

Variables used in the earnings forecast regressions, denoted with the subscript t, are winsorized at the 1st and 99th percentiles by year. We do so separately each year to account for changing distributions over time. This approach helps address measurement issues, like data errors or unusual transactions like major acquisitions or restructurings, without discarding possibly informative observations.

Our baseline sample also excludes firm-years for which we lack valid implied cost of capital and markup estimates. Invalid estimates include those missing input data, have negative implied cost of capital estimates, or are outside the top and bottom 1 percentiles.

While our dataset provides a large view of US corporate activity over several decades, some limitations merit acknowledgment. First, accounting practices have evolved over our sample period, potentially affecting the consistency of certain variables. For example, the treatment of goodwill, R&D expenses, and leases has changed significantly over time. Second, our reliance on as-reported financial data means we incorporate each firm's accounting choices without change, which may introduce some measurement variation across firms. Third, Compustat primarily covers public companies, so our findings may not generalize to private firms. This limitation is shared by De Loecker et al. (2020), who also focuses on publicly-traded companies. Finally, we note that the reporting of items like depreciation, interest, and accruals can vary across industries and over time, though we expect any resulting measurement errors to be random rather than systematic.

2.3 Earnings Forecast Estimation

We estimate firm-level earnings forecasts using a cross-sectional model that captures systematic patterns in future profitability. These forecasts serve as inputs for our implied cost of capital calculations, which in turn contribute to user cost and markup estimates. Hou et al. (2012) show cross-sectional earnings models outperform analyst forecasts for estimating implied cost of capital. Following this approach, we construct earnings predictions using accounting variables rather than analyst forecasts. It provides broader coverage and avoids potential analyst biases.

For each year between 1970 and 2022, we estimate the following pooled cross-sectional regression using the previous ten years of data:

$$E_{it+\tau} = \alpha_0 + \alpha_1 A_{it} + \alpha_2 D_{it} + \alpha_3 DD_{it} + \alpha_4 E_{it} + \alpha_5 NegE_{it} + \alpha_6 AC_{it} + \varepsilon_{it+\tau}$$

where $E_{it+\tau}$ denotes the earnings of firm i in year $t+\tau$ ($\tau = 1$ to 3), A_{it} is total assets, D_{it} is dividend payment, DD_{it} is a dummy variable that equals 1 for dividend payers and 0 otherwise, $NegE_{it}$ is a dummy variable that equals 1 for firms with negative earnings and 0 otherwise, and AC_{it} is accruals.

Each variable in our model serves a specific purpose. Total assets account for firm size effects, dividends signal financial stability, and current earnings capture persistence in performance. The negative earnings indicator addresses the different behavior of loss-making firms, while accruals help bridge the gap between accounting profits and cash flows.

We use a ten-year rolling window for estimation because it balances two competing needs: gathering enough observations for reliable coefficient estimates while adapting to changing economic conditions.

We obtain our earnings forecasts by multiplying the independent variables in year t with the coefficients from the cross-sectional regressions. For each firm-year observation, we forecast earnings up to three years ahead. To make sure our forecasts are out-of-sample, we use independent variables from the previous ten years of data only up to year t to forecast earnings for years $t+1$ through $t+3$.

This cross-sectional approach provides several advantages over firm-specific time-series models. It requires fewer observations per firm, making it applicable to companies with limited history. The rolling estimation window adapts to changing economic conditions by incorporating new information as it becomes available.

The models show strong predictive power, with average R-squared values exceeding 0.80 for one-year-ahead forecasts and remaining above 0.70 even for three-year projections. The coefficients demonstrate that earnings display significant persistence—past earnings substantially predict future earnings. The estimates also reveal that larger firms and dividend-paying companies generally deliver more predictable future performance.

Table 2: Coefficient Estimates of the Earnings Forecast Model

		Intercept	A_t	D_t	DD_t	E_t	$NegE_t$	AC_t	Adj.R^2
E_{t+1}	Coefficient	-1.0186	0.0041	0.2357	0.1593	0.7864	-0.0781	-0.1060	0.82
	t-statistic	-0.72	12.80	19.90	1.26	161.86	0.83	-28.31	
E_{t+2}	Coefficient	-0.2161	0.0088	0.4200	-0.0603	0.6562	-1.1098	-0.1376	0.75
	t-statistic	0.55	19.54	24.69	1.08	98.52	-0.29	-27.47	
E_{t+3}	Coefficient	0.8975	0.0125	0.5066	-0.2301	0.6176	-1.8266	-0.1452	0.71
	t-statistic	1.29	21.91	24.05	1.12	74.45	-0.66	-24.17	

Notes: This table presents the average coefficients from our annual cross-sectional earnings regressions. We estimate the model each year from 1970 to 2022 using the previous ten years of data. The dependent variable is future earnings for horizons of one to three years.

2.4 Implied Cost of Capital Estimation

We calculate each firm's implied cost of capital (ICC) using a residual income model that connects current market values to expected future cash flows (Gebhardt et al., 2001). This approach extends beyond our earnings forecasts to estimate the discount rate that investors apply to a firm's long-term income stream. Numerous studies validate the ICC approach, showing it predicts future returns and captures time-varying risk premia (Pástor et al., 2008; Lee et al., 2021). The ICC represents the internal rate of return that equates the firm's current market price to the present value of its expected future cash flows.

Following Gebhardt et al. (2001), we solve for the discount rate R in the equation:

$$M_t = B_t + \sum_{\kappa=1}^{11} \frac{E_t\left[(ROE_{t+\kappa} - R) \times B_{t+\kappa-1}\right]}{(1+R)^{\kappa}} + \frac{E_t\left[(ROE_{t+12} - R) \times B_{t+11}\right]}{R\,(1+R)^{11}}$$

where M_t is the market equity in year t, R is the implied cost of capital, B_t is the book equity, $E_t[\cdot]$ denotes market expectations based on information available at year t. $(ROE_{t+\kappa} - R) \times B_{t+\kappa-1}$ is the residual income in year $t+k$, defined as the difference between the return on book equity and the ICC multiplied by the book equity in the previous year.

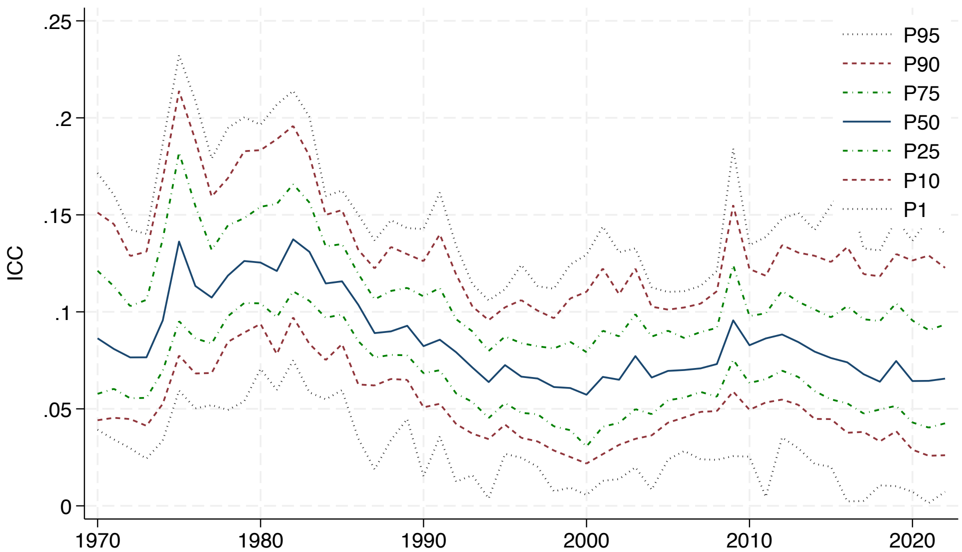
We estimate expected ROE in years $t+1$ to $t+3$ using earnings forecasts from the cross-sectional model and book equity values computed based on clean surplus accounting $B_{t+k} = B_{t+k-1} + E_{t+k} - D_{t+k}$, where E_{t+k} and D_{t+k} denote earnings and dividends in year $t+k$, respectively. Dividends are calculated using the current dividend payout ratio for firms with positive earnings. For firms with negative earnings, we use current dividends divided by $0.06\times$ total assets as an estimate of the payout ratio.

For horizons beyond our three-year model estimates, the expected ROE is assumed to mean-revert to the historical industry median value by year $t+11$. Mean reversion is achieved through simple linear interpolation between period $t+3$ ROE and the industry median ROE. After $t+11$, the residual income becomes a perpetuity.

This method gradually transitions firm-specific projections toward industry-median profitability levels. This mean reversion reflects the economic reality that competition tends to eliminate abnormal returns over time. High-performing firms attract competitors, while under-performing firms either improve or exit the market. We compute the industry target ROE as a moving median using the past ten years' ROE from all firms in the same Fama-French 49 industry. Following Gebhardt et al. (2001), we exclude loss firms when calculating the industry median ROE since persistent losses are not representative of long-term industry profitability.

Figure 1 shows the evolution of estimated ICCs at different percentiles over time. We weight ICC estimates by common equity value (CEQ).

Figure 1: Trends in Estimated ICCs (Equity Value Weighted)



2.5 Markup Estimation

We now turn to the estimation of firm-level markups, which measure the gap between the price firms charge and the marginal cost they incur. Our approach builds on Hall (1988), who first showed how to identify markups from production data, but we extend this method by incorporating financial market information to measure capital costs. When firms use labor and capital inputs with constant returns to scale technology, each input's cost share equals its output elasticity. The markup arises by comparing a firm's observed revenues to its total economic cost.

We treat labor costs as operating expenses (WL) and capital costs as the user cost of capital (R). The R combines the depreciation rate, the effective tax rate, and the weighted average cost of capital (WACC), which itself incorporates the firm's implied cost of equity (ICC) and cost of debt. This approach captures all relevant costs of capital—opportunity costs, physical depreciation, and tax effects. Under CRS, capital's share in total costs is the ratio of capital expenses to overall production expenses, and labor's share follows from operating expenses. Because these shares must sum to one, we infer that labor's output elasticity equals its share, and capital's elasticity equals its own share.

A firm's total economic cost equals:

$$\mathcal{C}_{it} = WL_{it} + \left(\mathrm{R}_{it} \times \mathrm{K}_{it}\right),$$

where K_{it} is invested capital, and R_{it} is the per-dollar cost of capital.

Our timing convention follows the "time-to-build" principle from capital theory. For production in year t (generating sales in year t), we use the capital stock and user cost from year $t-1$. This reflects two key economic realities. First, capital investment requires time to become productive—firms must order, install, and integrate new capital before it generates output. Second, the cost of capital should reflect the ex ante required return when investment decisions were made, not ex post realizations. Using lagged values ensures we measure the opportunity cost investors faced when committing capital, avoiding look-ahead bias that would arise from using contemporaneous market values. This timing convention is particularly important for our ICC estimates, as using current-year equity values to estimate the cost of capital for current-year production would introduce mechanical correlations between market valuations and measured markups. By using R_{it} and K_{it} measured as of the end of year $t-1$, we maintain the proper economic interpretation where markups reflect the wedge between prices and the true economic costs incurred to produce output.

If labor claims a fraction $\theta_{\ell,it}$ of total costs, then capital receives the remaining fraction, $\theta_{k,it} = 1 - \theta_{\ell,it}$. We interpret these fractions as the respective elasticities in the production function under CRS.

A firm's markup μ_{it} is the ratio of its price to its marginal cost. Because price times output equals total revenue, we can write:

$$\mu_{it} = \frac{\text{price} \times \text{output}}{\text{marginal cost} \times \text{output}}$$

We approximate marginal cost by applying the cost shares to the firm's operating expenses and user cost of capital. Our approach follows the production-based techniques as in De Loecker et al. (2020), but addresses measurement challenges emphasized by Foster et al. (2022) when inferring markups from firm-level data. Our main formula (ignoring intermediate inputs) becomes:

$$\mu_{it} = \frac{\text{Sales}_{it}}{WL_{it}} \times \left(1 - \frac{\mathrm{R}_{it}\,\mathrm{K}_{it}}{WL_{it} + \mathrm{R}_{it}\,\mathrm{K}_{it}}\right) = \frac{\text{Sales}_{it}}{WL_{it} + \mathrm{R}_{it}\,\mathrm{K}_{it}}. \tag{3}$$

The cost share approach offers an important advantage in its robustness to input classification. Because cost shares must sum to one under constant returns to scale, the method remains valid whether SG&A or R&D is treated as part of operating expenses or separated into distinct categories. As long as all productive inputs are included in the accounting, the total cost in the denominator captures the full economic cost of production. This property makes our estimates less sensitive to accounting conventions than alternative approaches that require precise specification of the production function (Traina, 2018; De Loecker et al., 2020).

In practice, deviations from perfect competition in the input market or CRS can affect these relationships. If a firm exhibits increasing returns to scale or secures labor at below-market rates, the measured markup may deviate from the actual gap between price and marginal cost. Nevertheless, empirical work Basu and Fernald (1997) suggests CRS is a reasonable approximation for many industries, and Syverson (2004) offers micro-data support.

This approach offers several advantages over traditional markup estimation methods. First, it uses readily available accounting data rather than requiring detailed product-level information. Second, it accommodates differences in capital intensity across firms and industries. Third, it accounts for variation in the cost of capital, recognizing that riskier firms face higher financing costs. Finally, it provides a theoretically grounded framework that connects firm-level decisions about investment and pricing to their financial market outcomes.

Once we have firm-level markups, we examine how they vary over time and across sectors. We also study whether markups differ for firms with dissimilar user costs of capital or different degrees of product-market concentration. Although the markup itself is unobservable, these production-based cost shares offer a systematic way to estimate how much prices exceed marginal costs. By incorporating our firm-specific ICC measures, we allow for the possibility that capital is more expensive for some firms than for others, introducing cross-sectional variation in markups within the same industry.

Table 3 presents our estimated variables. The user cost of capital averages 21% annually, combining a mean implied cost of capital of 10%, depreciation rate of 7%, and tax rate of 4%. The distribution shows considerable right skew, with the 99th percentile reaching 53%—more than double the mean.

Table 3: Summary Statistics of Estimated Variables

	N	Mean	SD	p1	p10	p25	Median	p75	p90	p99
Tax rate	160,996	0.04	0.06	-0.10	-0.01	0.00	0.03	0.07	0.11	0.23
Depreciation rate	160,996	0.07	0.10	0.00	0.02	0.03	0.05	0.08	0.12	0.27
Cost of debt	160,996	0.09	0.11	0.00	0.00	0.04	0.07	0.11	0.15	0.62
Implied cost of capital	160,996	0.10	0.06	0.00	0.04	0.06	0.09	0.13	0.17	0.30
WACC	160,996	0.10	0.14	0.01	0.04	0.07	0.09	0.13	0.17	0.30
User cost of capital	160,996	0.21	0.20	0.02	0.10	0.14	0.20	0.26	0.32	0.53
Markup	160,996	0.97	0.18	0.22	0.80	0.94	1.00	1.04	1.10	1.32
Output elasticity of capital	160,996	0.14	0.13	0.01	0.04	0.07	0.11	0.17	0.28	0.67
Variable cost share	160,996	0.86	0.13	0.33	0.72	0.83	0.89	0.93	0.96	0.99
Variable revenue share	160,996	0.96	0.60	0.37	0.73	0.83	0.90	0.95	1.04	3.34

Notes: This table presents summary statistics for all variables estimated using input variables described in Table 1.

Markups center near unity, with both mean (0.97) and median (1.00) indicating competitive conditions, though the 99th percentile reaches 1.32. The output elasticity of capital averages 14%. We observe notable differences across cost measures: while the cost of debt shows wide dispersion, the implied cost of capital distribution is more compressed.

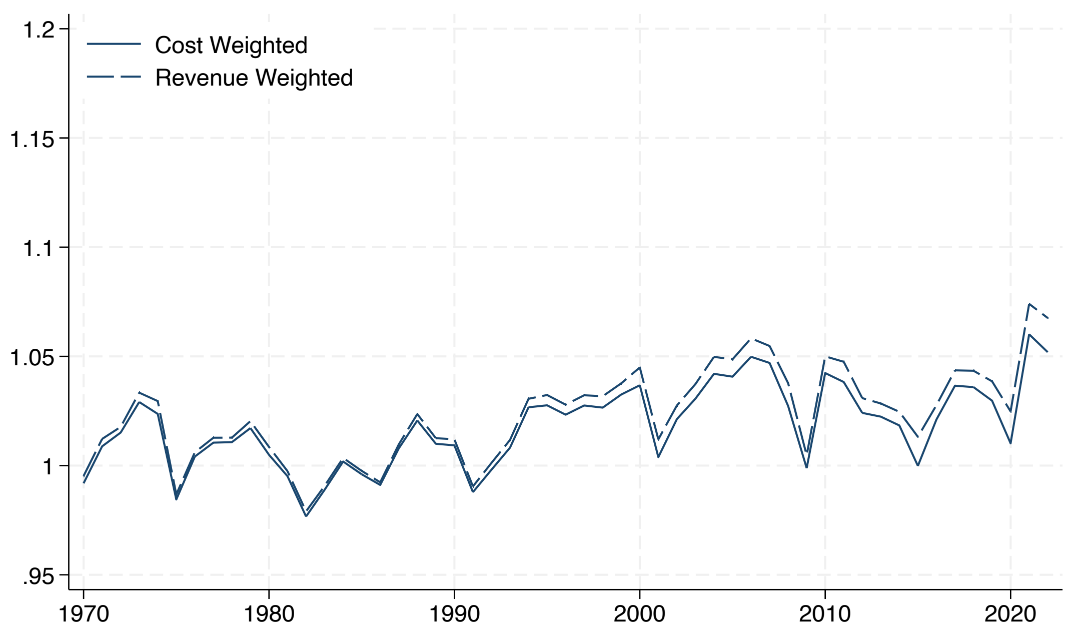
3 Empirical Findings

This section examines the distributional patterns of estimated firm-level markups from 1970 to 2022, focusing on their evolution over time, sectoral differences, and increasing dispersion. We also examine how markup trends diverge across industries, explore changes in the distribution's shape, and consider how firm-specific financing conditions help explain markup variation.

3.1 Overall Trends and Distributional Patterns

Figure 2 shows average markups from 1970 to 2022. We compute average markups as $\mu_t = \sum_i w_{it}\mu_{it}$, where w_{it} denotes firm-specific weights. We show markups computed using both cost-weighted and revenue-weighted approaches. The cost-weighted series assigns weights based on firm i's share of total variable costs in a given year, while the revenue-weighted series uses firm i's sales. The two series track closely over time, suggesting that the observed trends in markups are not driven by weighting choices but reflect underlying economic shifts.

Figure 2: Average Markup Estimates Over Time



This figure shows average markups from 1970 to 2022 computed as $\mu_t = \sum_i w_{it}\mu_{it}$, where w_{it} denotes firm-specific weights. The cost-weighted series (solid line) assigns weights based on firm i's share of total variable costs, while the revenue-weighted series (dashed line) uses firm i's sales share. Both series remain close to 1.0, indicating broadly competitive pricing throughout the sample period.

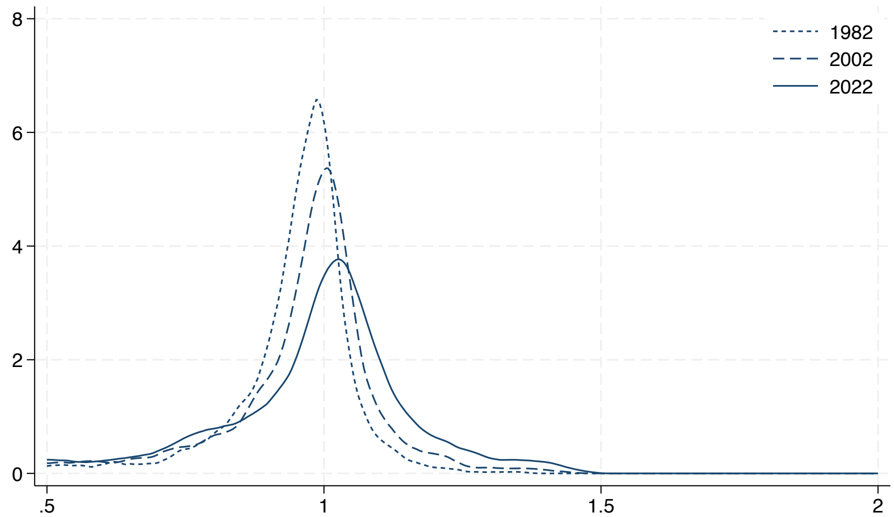
Average markups remained relatively stable until the early 1980s, followed by a period of increase through the 1990s and early 2000s. This rise coincides with broader shifts in market structure, technological change, and firm heterogeneity. These patterns align with the 'superstar firm' hypothesis of Autor et al. (2020) and the role of intangibles documented by Crouzet and Eberly (2023). The fluctuations suggest cyclical variation, likely reflecting economic shocks such as the financial crisis or dot-com bubble and subsequent recovery.

Our aggregation approach emphasizes cost weights rather than revenue weights, following Edmond et al. (2023). While conceptually the appropriate weight to map micro markups to a macro concept depends on the underlying model, commonly used CES aggregators largely imply some variant of cost weights. Even in the Edmond et al. (2023) framework, cost weights would be strictly correct only with no differences in output elasticities—contrary to our approach—but we adopt this weighting to keep things simple and comparable across studies. Our estimates are gross output markups rather than value-added markups. As Basu (2019) highlights, the implied value-added markups are often twice as high or more, the degree of which depends on the model and structure of production (Baqaee and Farhi, 2020).

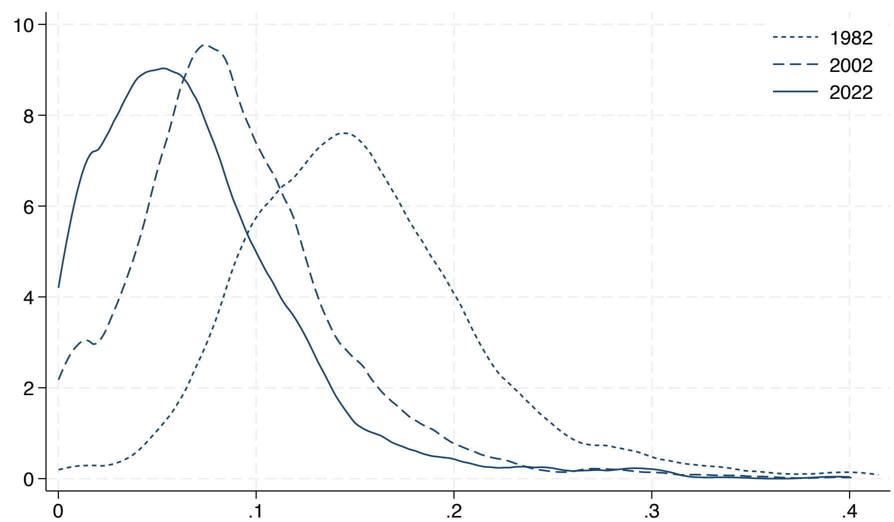
These stable averages mask important distributional changes. Panel (a) of Figure 3 presents kernel density estimates of unweighted markups for the years 1982, 2002, and 2022. The distribution has widened considerably with a more pronounced right tail in 2022. While many firms continue to operate with markups near 1, a growing subset of firms maintains substantially higher markups. This pattern indicates that markup growth is concentrated in a select group of firms rather than representing an economy-wide trend.

Panel (b) of Figure 3 presents kernel density estimates of unweighted implied costs of capital for the years 1982, 2002, and 2022. The distribution has shifted leftward over time. This is consistent with both the overall decrease in cost of equity over time (Duarte and Rosa, 2015) seen in Figure 1, and the decline in the number of public firms, especially smaller ones.

Figure 3: Kernel Density (unweighted)



(a): Markups



(b): Implied Cost of Capital

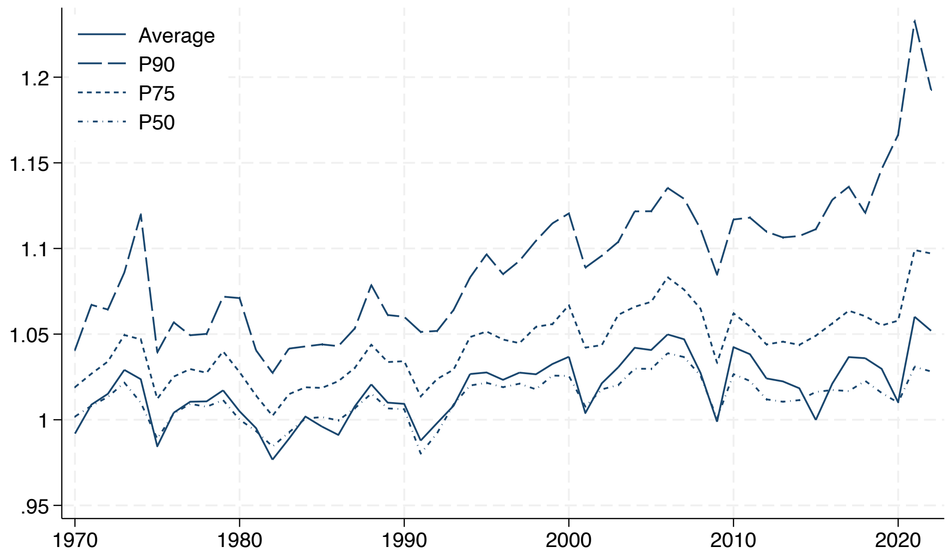
The widening markup distribution combined with compressing capital costs suggests an important connection. To investigate these distributional changes more precisely, we analyze different moments of the markup distribution while accounting for cost-weighting. Specifically, we compute percentile estimates for each year by ranking firm-level markups weighted by their share of total variable costs. This approach ensures that the percentiles are directly comparable to the weighted average markups reported earlier (De Loecker et al., 2020).

Figure 4 presents the trends in cost-weighted markups across different percentiles. Panel (a) tracks the 50th, 75th, and 90th percentiles, while Panel (b) focuses on the 90th, 95th, and 99th percentiles.

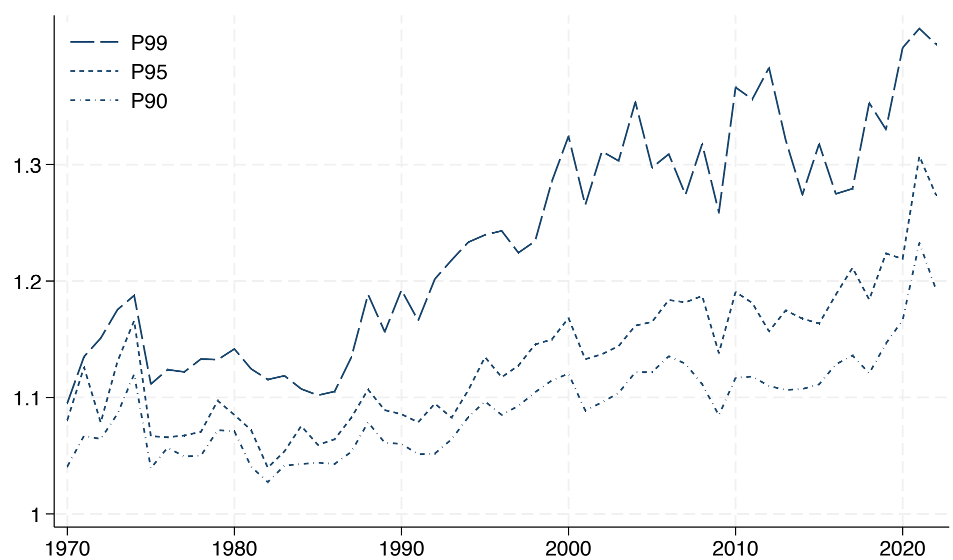
The increase in average markups is primarily driven by firms in the upper percentiles, particularly the top 10%. While the median markup (50th percentile) has remained relatively stable since 1970, the 90th percentile and above exhibit a sustained rise, with the 99th percentile showing the most pronounced increase.

This pattern reinforces the key finding that markup growth is not broad-based but rather concentrated among a subset of firms with the highest price-cost margins. The increasing dispersion in markups—evident from the widening gap between the median and upper percentiles—suggests that market power, technological advantages, or capital cost differentials may be enabling a select group of firms to command substantially higher markups over time.

Figure 4: Markup Distributions (cost-weighted)



(a): 90th, 75th, 50th percentiles



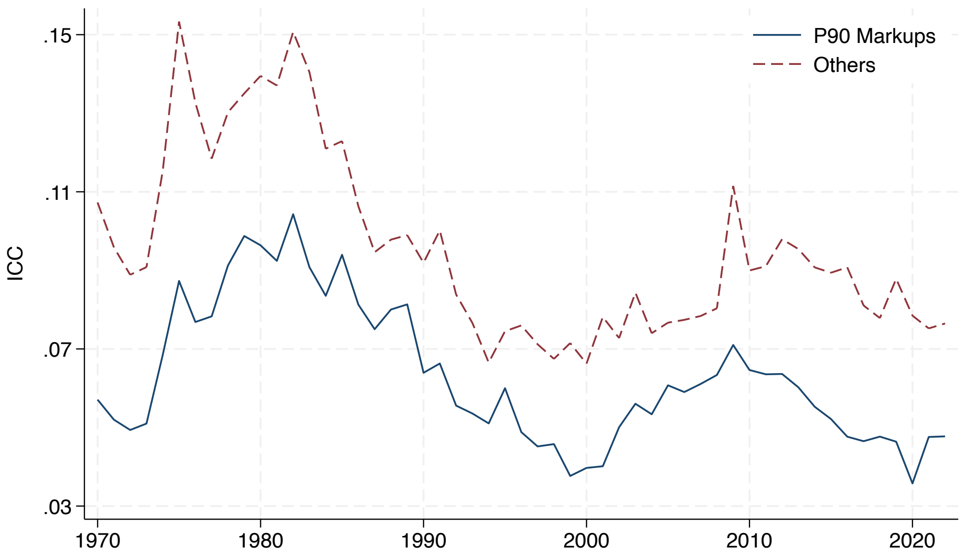
(b): 99th, 95th, 90th percentiles

3.2 Superstars and Costs of Capital

The growing dispersion in markups raises an important question: what drives the divergence between high- and low-markup firms? Our theoretical framework suggests that differences in the cost of capital could be a potential mechanism. To test this hypothesis, Figure 5 compares the average ICC for firms in the top 10% of the markup distribution (P90 markups) to that of all other firms.

We weight ICCs by equity value. The figure reveals a strong negative relationship between ICC and markups: firms with the highest markups consistently exhibit lower financing costs compared to the broader firm population. While ICCs for the top 10% of markup firms have largely remained below 7% since 1990, the ICC for all other firms has been substantially higher, often exceeding 10%.

Figure 5: Average ICCs for the Top 10% vs. All Others



This pattern reveals an important economic mechanism: firms with better financing conditions—whether due to lower risk premia or stronger capital market access—are able to sustain higher markups. The persistent gap between these groups suggests that financing advantages create durable competitive positions, consistent with the superstar firm dynamics described in Liu et al. (2022). Table 4 quantifies this relationship more precisely. In particular, Column (5) of Table 4 demonstrates that firm-year variations in ICC explain nearly 60% of the variation in markups (adjusted R^2=0.59).

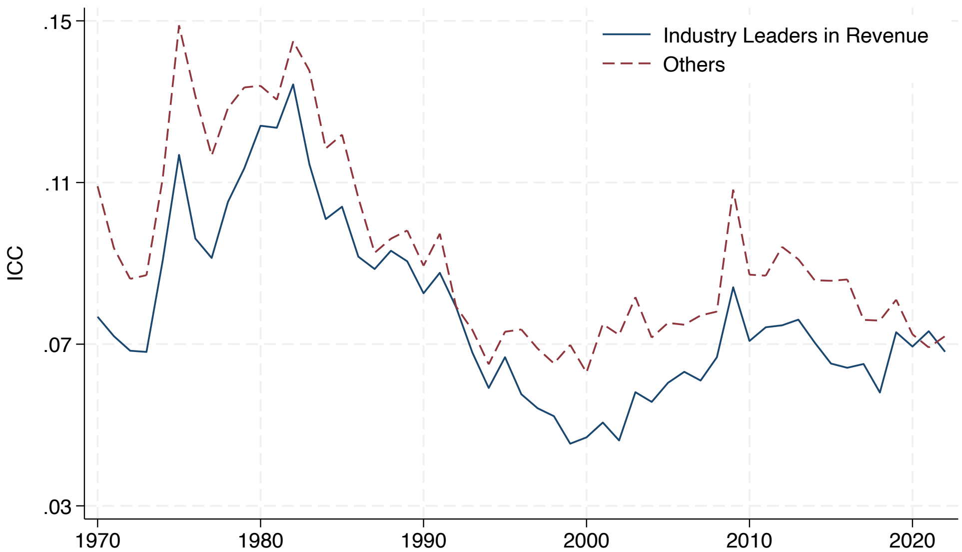
Table 4: Regressions of Markup on ICC

	(1)	(2)	(3)	(4)	(5)
Dependent Variable: Ln(Markup)					
Implied cost of capital	0.000	-0.001***	-0.003***	-0.005***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Year FE	N	Y	Y	Y	N
Industry FE	N	N	Y	N	N
Firm FE	N	N	N	Y	Y
Industry × Year FE	N	N	N	N	Y
Clusters (Firms)	17,053	17,053	16,731	14,488	14,232
Adjusted R-Squared	-0.000	0.017	0.091	0.570	0.596
Observations	160,996	160,996	158,478	158,431	155,978

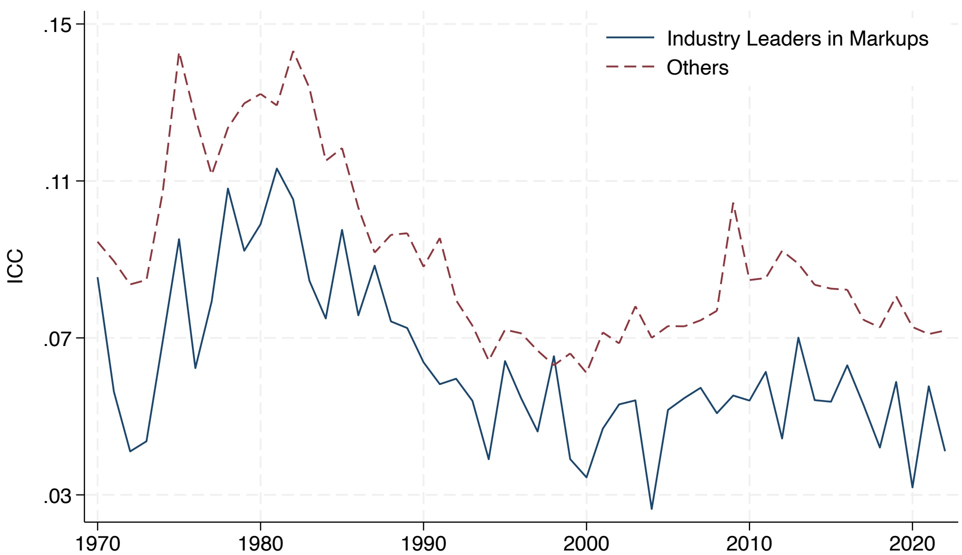
Notes: This table presents OLS regressions of the form: $\ln(\mu_{it}) = \alpha + \beta \cdot ICC_{it} + \delta_i + \delta_t + \epsilon_{it}$, where μ_{it} is the markup for firm i in year t, and ICC_{it} is the implied cost of capital. Column (1) shows the univariate relationship. Column (2) adds year fixed effects. Column (3) adds both year and industry fixed effects. Column (4) includes both year and firm fixed effects. Column (5) includes firm and industry-year fixed effects. The sample spans 1970 to 2022. Standard errors clustered at the firm-level are reported in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

The strong negative relationship between markups and financing costs raises an important question: are these advantages specific to high-markup firms, or do they extend to large firms more generally? In Figure 6 Panel (a) and (b), we plot the average implied costs of capital for industry leaders in sales and markups, respectively. While industry leaders in markups have clearly lower ICCs relative to others, the corresponding gap is narrower between industry leaders in revenue and others.

Figure 6: Average ICCs of Industry Leaders vs. Others



(a): Leaders in Revenue

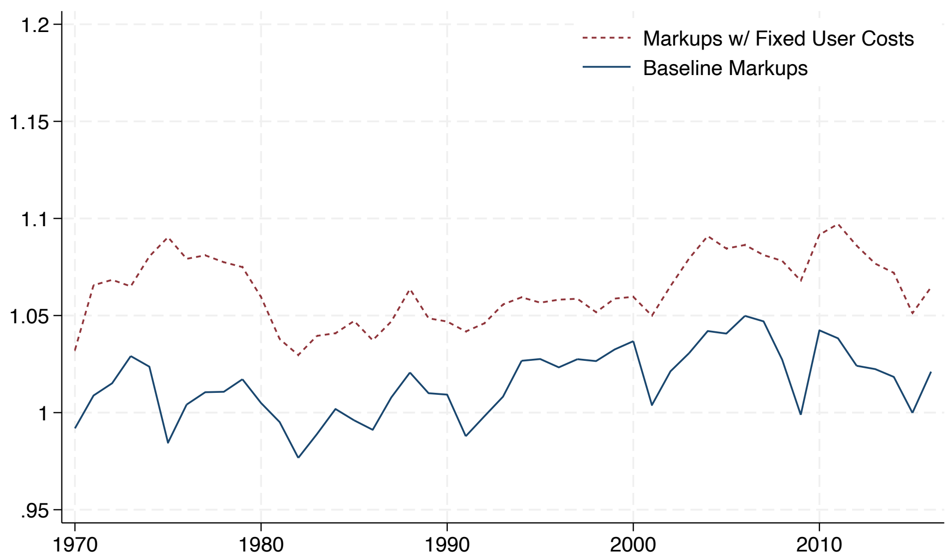


(b): Leaders in Markups

The sharper ICC gap for markup leaders versus revenue leaders suggests that financing advantages are more closely tied to pricing power than to size alone. This distinction helps explain why some large firms maintain competitive pricing while others command substantial markups. To quantify the aggregate importance of these firm-specific financing costs, we analyze how markups vary under different ICC assumptions. Figure 7 compares our *baseline cost-weighted markup estimates*, which incorporate firm-year ICCs, to alternative scenarios where user costs of capital are held constant at an aggregate level. The solid line represents our main estimates, while the dashed lines show markups calculated using fixed user cost of capital values.

Ignoring firm-level ICC heterogeneity systematically inflates markup estimates. When a constant user cost of capital is used, markup estimates are consistently higher and smoother, failing to capture the extent of variability seen in the baseline estimates. This suggests that a significant portion of the observed variation in markups is attributable to differences in firms' user costs of capital and financing costs, reinforcing the central role of capital markets in explaining rising markup dispersion.

Figure 7: Comparison of Markup Estimates: ICC-Adjusted vs Fixed User Cost



This figure compares markup estimates under different user cost assumptions. The solid line shows our baseline estimates using firm-specific implied cost of capital (ICC). The dashed line shows markups calculated using a fixed user cost of capital for all firms (similar to the De Loecker et al. approach). ICC-adjusted markups are approximately 15% lower and show greater variation. Ignoring firm-level heterogeneity in financing costs systematically overstates aggregate markups.

To assess the persistence of firm-specific capital costs and pricing behavior, we regress ICCs and markups on their lagged values. Table 5 presents the persistence estimates for ICC, while Table 6 reports the corresponding results for markups. Across all specifications, we find strong serial correlation in both variables, suggesting that firms' financing conditions and pricing power exhibit substantial persistence over time.

Table 5: Persistence - Cost of Capital

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: ICC								
ICC_{t-1}	0.687***	0.614***	0.597***	0.336***	0.507***	0.454***	0.450***	0.297***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.007)	(0.007)	(0.007)	(0.008)
ICC_{t-2}					0.156***	0.149***	0.148***	0.079***
					(0.006)	(0.006)	(0.006)	(0.006)
ICC_{t-3}					0.088***	0.084***	0.083***	0.036***
					(0.006)	(0.006)	(0.006)	(0.006)
ICC_{t-4}					0.047***	0.039***	0.036***	-0.001
					(0.006)	(0.006)	(0.006)	(0.006)
ICC_{t-5}					0.033***	0.044***	0.041***	-0.013**
					(0.005)	(0.005)	(0.005)	(0.006)
Year FE	N	Y	Y	Y	N	Y	Y	Y
Industry FE	N	N	Y	N	N	N	Y	N
Firm FE	N	N	N	Y	N	N	N	Y
Clusters (Firms)	14,132	14,132	13,884	12,142	7,859	7,859	7,746	6,893
Adjusted R-Squared	0.471	0.529	0.534	0.599	0.556	0.594	0.596	0.634
Observations	135,819	135,819	133,775	133,829	78,916	78,916	77,806	77,950

Notes: This table reports OLS regressions of the form: $ICC_{it} = \alpha + \sum_{k=1}^{L} \beta_k \cdot ICC_{i,t-k} + \delta_i + \delta_t + \epsilon_{it}$, where ICC_{it} is the implied cost of capital for firm i in year t, and L is the number of lags (ranging from 1 to 5). Columns differ in the inclusion of fixed effects and the number of lags. The coefficients measure the persistence of cost of capital over time. The sample spans 1970 to 2022. Standard errors clustered at the firm-level are reported in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

For ICCs, the one-year lag coefficient remains above 0.29 across all models, with significance levels confirming the robustness of the estimates. Even at longer lags, ICCs display notable persistence, with coefficients remaining positive and statistically significant up to three years in most specifications. This indicates that firms with higher capital costs tend to experience persistently elevated financing expenses, potentially constraining their long-term investment and growth.

Table 6: Persistence - Markups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: Ln(Markup)								
$\ln(\text{Markup}_{t-1})$	0.780***	0.781***	0.758***	0.429***	0.580***	0.577***	0.571***	0.367***
	(0.007)	(0.007)	(0.008)	(0.012)	(0.015)	(0.015)	(0.016)	(0.017)
$\ln(\text{Markup}_{t-2})$					0.116***	0.124***	0.120***	0.047***
					(0.016)	(0.016)	(0.016)	(0.016)
$\ln(\text{Markup}_{t-3})$					0.095***	0.094***	0.091***	0.055***
					(0.014)	(0.014)	(0.015)	(0.013)
$\ln(\text{Markup}_{t-4})$					0.082***	0.081***	0.080***	0.041***
					(0.013)	(0.014)	(0.014)	(0.012)
$\ln(\text{Markup}_{t-5})$					0.028***	0.028**	0.027**	-0.007
					(0.011)	(0.011)	(0.011)	(0.011)
Year FE	N	Y	Y	Y	N	Y	Y	Y
Industry FE	N	N	Y	N	N	N	Y	N
Firm FE	N	N	N	Y	N	N	N	Y
Clusters (Firms)	14,132	14,132	13,884	12,142	7,859	7,859	7,746	6,893
Adjusted R-Squared	0.519	0.529	0.534	0.619	0.499	0.511	0.512	0.588
Observations	135,819	135,819	133,775	133,829	78,916	78,916	77,806	77,950

Notes: This table reports OLS regressions of the form: $\mu_{it} = \alpha + \sum_{k=1}^{L} \beta_k \cdot \mu_{i,t-k} + \delta_i + \delta_t + \epsilon_{it}$, where μ_{it} is the markup for firm i in year t, and L is the number of lags (ranging from 1 to 5). Columns differ in the inclusion of fixed effects and the number of lags. The coefficients measure the persistence of markups over time. The sample spans 1970 to 2022. Standard errors clustered at the firm-level are reported in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

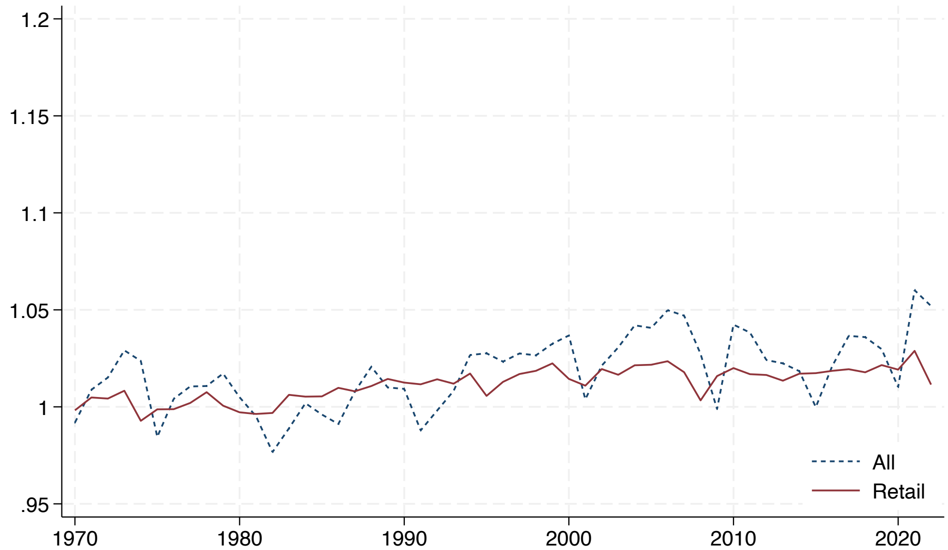
Similarly, markups exhibit substantial persistence, with an autoregressive coefficient near 0.58 for the one-year lag in models without firm fixed effects. When controlling for firm-specific heterogeneity, the one-year lag coefficient remains significant but decreases to approximately 0.36, indicating that while firm-specific factors contribute to markup stability, there is also considerable variation across firms. Higher-order lags show a gradual decline in explanatory power, but remain statistically significant, reinforcing the idea that firms with high markups tend to maintain pricing power over extended periods. These results provide further evidence that differences in financing conditions and firm-specific characteristics play a key role in markup dynamics.

3.3 Production Technology and Intangibles

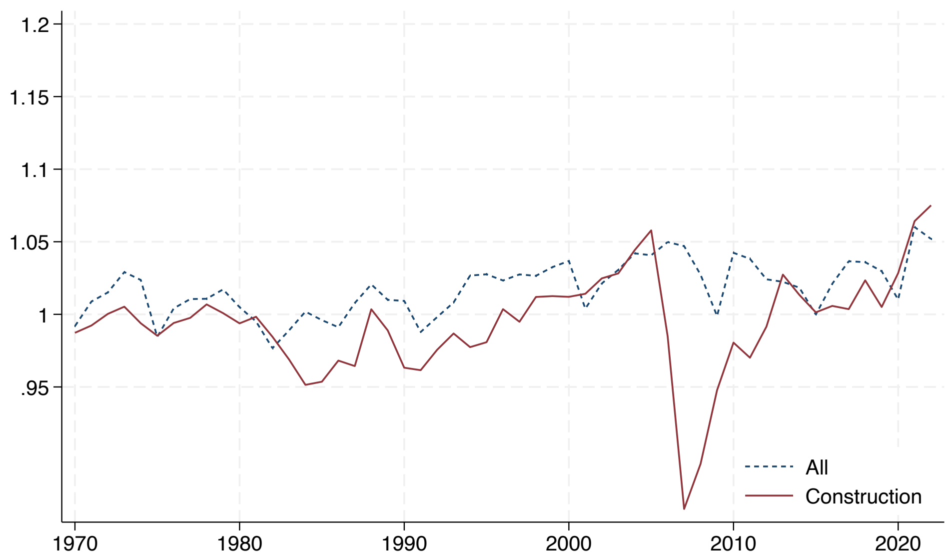
In this section, we examine how production technology and intangibles are related to documented markup trends.

Figures 8 and 9 present cost-weighted average markups by sector, highlighting distinct patterns across industries with different characteristics. We use the Fama-French 49 industry classifications to categorize firms into different industries.

Figure 8: Markups Trends by Industry - Competitive



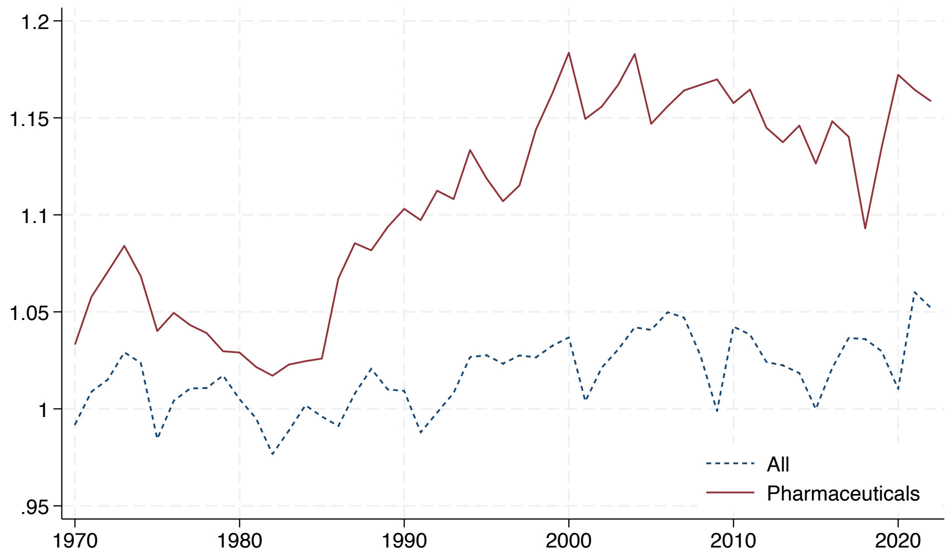
(a): Retail



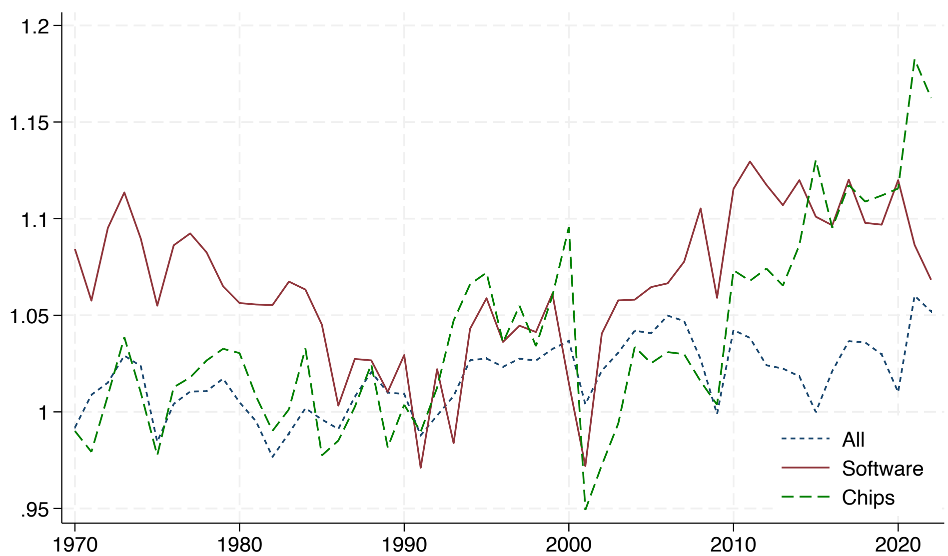
(b): Construction

In Figure 8 we examine markup trends of highly competitive industries, specifically retail and construction. Markups in retail largely track the overall trend, with a flatter trajectory and minimal volatility, suggesting sustained competitive pressures that limit firms' ability to price above marginal cost. The markup trends for construction shown in Panel (b) also remain similar to the overall trend and closer to 1. However, markups show a notable drop around the financial crisis in 2008-9, where construction firms were directly affected.

Figure 9: Markups Trends by Industry - Tech



(a): Pharmaceuticals



(b): Software & Chips

In Figure 9 we examine markup trends of technology focused industries, specifically pharmaceuticals, software, and chips. These industries, which tend to be highly R&D and intangible intensive firms, exhibit both higher average markups and greater variability over time. Markup trends for software and chips plotted in Panel (b) again show a sharp drop around the dot-com bubble burst. Notably, markups in these sectors notably increase over our sample period, consistent with the rise of high-margin firms in software and digital platforms.

A takeaway is that markup increases are not economy-wide but rather concentrated in industries where firm-level heterogeneity in technology and financing might play a larger role. We conduct regression analyses to further investigate this relationship.

Table 7: Regressions of Markup on Intangibles

	Intangibles = 1-PPENT/ICAPT				Intangibles = XRD/XOPR			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Markup)								
Intangible Assets Intensity	-0.002	-0.004	-0.002	-0.001				
	(0.002)	(0.003)	(0.001)	(0.001)				
R&D Intensity					-0.591***	-0.680***	-0.346***	-0.351***
					(0.055)	(0.087)	(0.116)	(0.117)
Ln(Assets)				0.031***				0.032***
				(0.002)				(0.002)
Year FE	N	Y	Y	Y	N	Y	Y	Y
Industry FE	N	Y	N	N	N	Y	N	N
Firm FE	N	N	Y	Y	N	N	Y	Y
Clusters (Firms)	14,126	13,878	12,138	12,138	14,132	13,884	12,142	12,142
Adjusted R-Squared	0.000	0.077	0.546	0.552	0.059	0.129	0.549	0.555
Observations	135,726	133,694	133,738	133,738	135,819	133,775	133,829	133,829

Notes: This table presents OLS regressions of the form: $\mu_{it} = \alpha + \beta_1 \cdot \text{Intangible Intensity}_{it} + \beta_2 \cdot \text{R\&D Intensity}_{it} + \gamma \cdot \ln(Assets_{it}) + \delta_i + \delta_t + \epsilon_{it}$. Intangible intensity is measured as intangible assets divided by total assets. R&D intensity is R&D expenses divided by operating expenses. Columns 1-4 examine intangible intensity with varying fixed effects. Columns 5-8 examine R&D intensity. The sample spans 1970 to 2022. Standard errors clustered at the firm-level are reported in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels, respectively.

The results illustrate an absence of a statistically significant association between intangible asset intensity and markups (Columns 1-4). In contrast, R&D intensity exhibits a strong negative relationship with markups (Columns 5-8), with coefficients ranging from -0.35 to -0.59. This suggests that firms with higher R&D spending relative to operating expenses tend to have lower markups, possibly due to the high upfront costs of innovation or competitive pressures in R&D-intensive industries.

Controlling for firm size (Columns 4 and 8) shows that larger firms tend to have higher markups, as indicated by the positive and significant coefficient on ln(Assets). This aligns with the broader literature on firm size and market power.

Overall, these findings highlight the heterogeneity in the relationship between intangible investments and markups and the distinction between capitalized intangible assets and ongoing R&D expenditures.

4 Concluding Remarks

We combine two previously separate methodologies—cost share markup estimation and implied cost of capital approaches as in Gebhardt et al. (2001) and Hou et al. (2012)—to measure markups using financial statement data. This combination fills an important gap, connecting production and finance approaches that previously existed separately despite their natural fit. Accounting for firm-specific financing costs shows average markups have remained stable at around 1.0 since 1980, not rising as previous studies suggest. The real story is increasing markup dispersion: firms in the top quartile increased markups from 1.03 to 1.14, while bottom-quartile firms saw markups fall from 0.9 to 0.8. We find firms with low capital costs maintain markups higher than typical firms. This negative relationship between financing costs and markups provides a missing mechanism in the superstar firm literature—cheaper access to capital might itself generate pricing power.

Our approach solves persistent measurement problems in the production approach to markups. The Hall (1988) cost share method requires accurate estimation of all input costs, but previous implementations used uniform or simplified capital costs. By incorporating firm-specific financing costs derived from market data, we allow capital costs to vary across firms and time. This variation matters—using simplified capital costs systematically distorts markup estimates. Our method directly addresses concerns raised by Basu (2019) and Syverson (2019) about markup measurement problems, bridging industrial organization and finance by showing how financial market conditions can inform our understanding of market power trends.

Our analysis faces two notable limitations. We focus on public firms due to data requirements for estimating implied cost of capital, which may not capture private firm dynamics. The cost share approach also relies on constant returns to scale.[3] Nevertheless, incorporating firm-specific variation in financing costs provides a more complete understanding of markup dynamics than previous approaches.

Future research should examine how companies invest in assets that don't appear on balance sheets, such as research and development expenses. These intangible investments create value but aren't treated as capital in standard accounting. This matters for our markup measures because a tech company might look like it has high markups when it's actually earning returns on its unmeasured intangible assets. Using our approach, researchers could directly measure how intangibles affect both capital costs and markups. Eisfeldt and Papanikolaou (2013) show organization capital carries unique risks affecting expected returns. Similarly, Crouzet and Eberly (2023) show how intangibles can generate both markups and higher user costs. This research direction is especially important as the economy shifts toward intangible-intensive industries.

Our findings raise practical questions for future research on competition. Do tech firms fund growth differently than manufacturing firms? Are rising profits a sign of market power or just returns on

[3] Proxy estimators in the style of Olley and Pakes (1996) largely require constant returns to scale to identify markups, as well (Flynn et al., 2019).

unmeasured investments? Why do some firms get better financing terms than others? Answering these questions would help explain whether rising profit gaps come from declining competition or technological change. This matters for policy debates about industry concentration and inequality. By showing how both production methods and financing costs affect our inference about market power, our paper offers economists new tools to tackle these important questions.

References

Abel, A. and J. Eberly (2011). How q and cash flow affect investment without frictions: An analytic explanation. *Review of Economic Studies 78*(4), 1179–1200.

Autor, D., D. Dorn, L. Katz, C. Patterson, and J. Van Reenen (2020). The fall of the labor share and the rise of superstar firms. *Quarterly Journal of Economics 135*(2), 645–709.

Ayyagari, M., A. Demirgüç-Kunt, and V. Maksimovic (2024). The rise of star firms: Intangible capital and competition. *Review of Financial Studies 37*(3), 882–949.

Balvers, R., L. Gu, and D. Huang (2017). Profitability, value, and stock returns in production-based asset pricing without frictions. *Journal of Money, Credit and Banking 49*(7), 1621–1651.

Baqaee, D. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *Quarterly Journal of Economics 135*(1), 105–163.

Barkai, S. (2020). Declining labor and capital shares. *Journal of Finance 75*(5), 2421–2463.

Basu, S. (2019). Are price-cost markups rising in the United States? a discussion of the evidence. *Journal of Economic Perspectives 33*(3), 3–22.

Basu, S. and J. Fernald (1997). Returns to scale in US production: Estimates and implications. *Journal of Political Economy 105*(2), 249–283.

Basu, S. and J. Fernald (2001). Why is productivity procyclical? Why do we care? In *New Developments in Productivity Analysis*, pp. 225–302. National Bureau of Economic Research.

Boar, C., D. Gorea, and V. Midrigan (2022). Why are returns to private business wealth so dispersed? Working paper, National Bureau of Economic Research.

Bond, S., A. Hashemi, G. Kaplan, and P. Zoch (2021). Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *Journal of Monetary Economics 121*, 1–14.

Bond, S. and M. Söderbom (2005). Adjustment costs and the identification of Cobb Douglas production functions. Working paper, Institute for Fiscal Studies.

Cochrane, J. (1991). Production-based asset pricing and the link between stock returns and economic fluctuations. *Journal of Finance 46*(1), 209–237.

Cochrane, J. (2009). *Asset pricing: Revised edition.* Princeton University Press.

Cooper, R. and J. Ejarque (2001). Exhuming q: Market power vs. capital market imperfections. Working Paper 8182, National Bureau of Economic Research.

Covarrubias, M., G. Gutiérrez, and T. Philippon (2020). From good to bad concentration? US industries over the past 30 years. *NBER Macroeconomics Annual 34*(1), 1–46.

Crouzet, N. and J. Eberly (2023). Rents and intangible capital: A q+ framework. *Journal of Finance 78*(4), 1873–1916.

Damodaran, A. (2007). Return on capital (ROC), return on invested capital (ROIC) and return on equity (ROE): Measurement and implications. Teaching note, New York University Stern School of Business.

David, J. and V. Venkateswaran (2019). The sources of capital misallocation. *American Economic Review 109*(7), 2531–2567.

Davis, C., A. Sollaci, and J. Traina (2024). Profit puzzles and the fall of public-firm profit rates. Working Paper 42, Kelley School of Business.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics 135*(2), 561–644.

Di Tella, S., C. Maglieri, and C. Tonetti (2024). Risk markups. Working paper, Social Science Research Network.

Duarte, F. and C. Rosa (2015). The equity risk premium: a review of models. *Economic Policy Review 21*(2), 39–57.

Edmond, C., V. Midrigan, and D. Y. Xu (2023). How costly are markups? *Journal of Political Economy 131*(7), 1619–1675.

Eisfeldt, A. and D. Papanikolaou (2013). Organization capital and the cross-section of expected returns. *Journal of Finance 68*(4), 1365–1406.

Fama, E. and K. French (1992). The cross-section of expected stock returns. *Journal of Finance 47*(2), 427–465.

Farhi, E. and F. Gourio (2018). Accounting for macro-finance trends: Market power, intangibles, and risk premia. *Brookings Papers on Economic Activity*, 147–223.

Flynn, Z., A. Gandhi, and J. Traina (2019). Measuring markups with production data. Working paper, Social Science Research Network.

Foster, L., J. Haltiwanger, and C. Tuttle (2022). Rising markups or changing technology? Working paper, National Bureau of Economic Research.

Frank, M. and T. Shen (2016). Investment and the weighted average cost of capital. *Journal of Financial Economics 119*(2), 300–315.

Gandhi, A., S. Navarro, and D. Rivers (2020). On the identification of gross output production functions. *Journal of Political Economy 128*(8), 2973–3016.

Gebhardt, W., C. Lee, and B. Swaminathan (2001). Toward an implied cost of capital. *Journal of Accounting Research 39*(1), 135–176.

Grullon, G., Y. Larkin, and R. Michaely (2019). Are US industries becoming more concentrated? *Review of Finance 23*(4), 697–743.

Hall, R. (1988). The relation between price and marginal cost in US industry. *Journal of Political Economy 96*(5), 921–947.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica 50*(1), 213–224.

Hou, K., M. van Dijk, and Y. Zhang (2012). The implied cost of capital: A new approach. *Journal of Accounting and Economics 53*(3), 504–526.

Jorgenson, D. (1963). Capital theory and investment behavior. *American Economic Review 53*(2), 247–259.

Kahle, K. and R. Stulz (2017). Is the US public corporation in trouble? *Journal of Economic Perspectives 31*(3), 67–88.

Karabarbounis, L. and B. Neiman (2019). Accounting for factorless income. *NBER Macroeconomics Annual 33*(1), 167–228.

Klette, T. J. and Z. Griliches (1996). The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics 11*(4), 343–361.

Lee, C., E. So, and C. Wang (2021). Evaluating firm-level expected-return proxies: implications for estimating treatment effects. *Review of Financial Studies 34*(4), 1907–1951.

Liu, E., A. Mian, and A. Sufi (2022). Low interest rates, market power, and productivity growth. *Econometrica 90*(1), 193–221.

Olley, S. and A. Pakes (1996). Market share, market value and innovation in a panel of british manufacturing firms. *Econometrica 64*(6), 1263–1297.

Pástor, L., M. Sinha, and B. Swaminathan (2008). Estimating the intertemporal risk–return tradeoff using the implied cost of capital. *Journal of Finance 63*(6), 2859–2897.

Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy 112*(6), 1181–1222.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature 49*(2), 326–365.

Syverson, C. (2019). Macroeconomics and market power: Context, implications, and open questions. *Journal of Economic Perspectives 33*(3), 23–43.

Traina, J. (2018). Is aggregate market power increasing? Production trends using financial statements. Working Paper 18, Stigler Center for the Study of the Economy and the State.

Van Reenen, J. (2018). Increasing differences between firms: Market power and the macro-economy. Discussion paper, Centre for Economic Performance.