Measuring Markups:

Revisiting the Cost Accounting Approach

Vincent Mastantuno*

Updated: December 15, 2024

Most recent version available here.

Abstract

Among the ways to measure market power, special attention has always been given to the wedge between marginal cost and price: the markup. This paper investigates the validity of firm-level markup estimation techniques, focusing on the cost accounting (CA) approach—a straightforward, transparent, and data-thrifty alternative that, despite its origins in the early industrial organization literature, has seen little modern application. While the production function (PF) approach has quickly become the workhorse method in macroeconomic applications, several high-profile criticisms have brought its validity into question. My findings show that the rise in U.S. markups is robust to CA, suggesting that the trend is not driven by PF's biases. In fact, variation in PF estimates is driven almost entirely by variation in its 'accounting component,' rather than its 'production function component,' which contains the well-known biases. For validation, I develop a novel test based on Dorfman-Steiner's (1954) advertising equation, concluding that the cost accounting approach has a higher signal-to-noise ratio, while both measures retain some signal of underlying markups. The data-thrifty nature of CA makes it feasible for a broader range of applications. To highlight this, I conclude with several examples that play into CA's strengths. Collectively, these results suggest that CA should be preferred to PF, although it can sometimes be useful to try both.

Keywords: Markups, Market Power, Cost Accounting, Advertising

JEL Codes: D2, D4, L1, L2, M3, M4

*University of Colorado Boulder, Vincent.Mastantuno@Colorado.edu. I would like to thank Miles Kimball, Alessandro Peri, and Scott Savage for their valuable advice. I am also grateful to Martin Boileau, Matthew Nagler, Murat Iyigun, Richard Mansfield, Adam McCloskey, Terra McKinnish, Jeffrey Zax, Tania Barham, Francisca Antman, Adam Pfander, Austin Kennedy, Aditya Rachmanto, Dongkyu Yang, Lucas Ladenburger, Sean Strunk, Alec Costanza, Paro Suh, and all members of the Macroeconomics, IO, Applied, and Graduate Student Brown bags at University of Colorado Boulder for their helpful comments. I would also like to thank the Leeds business school at the Univerity of Colorado, and WRDS hosted by the Wharton School of business for data access. All opinions expressed herein, and any remaining errors are my own.

1 Introduction

Competition–and its opposite, market power–have been central to economics since the field's inception. Among the ways to measure market power, special attention has always been given to the wedge between marginal cost and price, the markup. In perfect competition, firms compete directly and prices are driven down to marginal cost, leading to a zero markup (or a markup *ratio* of price over marginal cost equal to one). When facing less ferocious competition, firms can raise profits by setting prices above costs–a positive markup. Granular markup data, representative at a macro-economic level, are extremely valuable both as a means to answer important economic questions directly, and to test and calibrate many key models in economics.

This paper investigates the validity of firm-level markup estimation techniques. Of particular interest is the production function (PF) approach, which has been used to document a rise in aggregate markups in the United States. While PF has quickly become the workhorse method in macroeconomic applications, several high-profile criticisms have brought its validity into question. Despite its origins in the early industrial organization literature, the cost accounting (CA) approach–a straightforward, transparent, and data-thrifty alternative– has received relatively little attention. To my knowledge, this is the first paper to directly compare the performance of the two approaches to markup estimation, and the first systematic evaluation of the cost accounting approach of any sort.

My findings show that the rise in U.S. markups is robust to CA, suggesting that the trend is not driven by PF's biases. The production function estimates can be decomposed into an "accounting component" constructed directly from accounting data as the ratio of revenue to expenditures in one variable input, and an estimated output elasticity that is the source of the well-known biases of PF. I show that variation in PF estimates are driven almost entirely by its accounting component: the ratio of revenue to noncapital costs. While the literature has been primarily concerned with noise and bias in the output elasticity measure, that component simply does not contain much of a signal. Behind the complexity of PF, nearly identical results can be obtained with use of revenue to noncapital cost component. However, the ratio of revenue to noncapital costs by itself lacks a micro-foundation as a measure of the markup by itself. CA, on the other hand, can be micro-founded explicitly, and can be implemented directly from accounting data without the need to estimate production function parameters.

For validation, I develop a novel test based on the Dorfman and Steiner [1954] advertising equation, and find that both measures contain some signal of underlying markups, but CA outperforms PF in its signal-to-noise ratio. CA dramatically outperforms PF in the sector and time cross sections—the dimensions in which the production function component of PF is allowed to vary. Collectively these results suggest

both approaches give markups that have a substantial signal of true underlying markups, though they have different errors (and different error variances). I argue that on balance the evidence should motivate more widespread use of the cost accounting approach. And, indeed, because it is less data-intensive than the other two methods, there are contexts in which only the cost accounting method can be used. I focus on three applications that play into this strength of CA: a case study of the markups of tech giants, an investigation of markups over the firm life cycle, and an investigation of the cyclicality of markups in response to aggregate supply and demand shocks.

1.1 Background

Markups matter for central considerations of economics. Firm-level markup data representative of the macro-economy is invaluable both for answering important questions directly, and testing and calibrating economic models.

In static models, markups drive a wedge between marginal cost of production and marginal benefit of consumption, mechanically leading to inefficiency. Wedges of this sort erode the efficiency of general equilibrium in the Arrow-Debreu sense[1], moving the economy away from the optimality results of the first and second theorem of welfare economics.[2] As such, aggregate movements in and of themselves are an important measure of the static efficiency of the economy; policymakers and economists alike ought to be aware of their level and trend.

In the long run, the relationship between markups and efficiency is more complex. Pursuit of future profits can lead to innovation, so imperfect competition is integral to endogenous technological change.[3] Such models of endogenous technological change are founded on increasing returns to scale and monopolistic competition, with revenue generated from markups driving research and human capital investment—leading to knowledge spillovers and (under the right conditions) improvements in long-run dynamic efficiency. Known since the dawn of endogenous growth theory, the centrality of markups features prominently in recent growth theory.[4] Markups are therefore at the core of the conflict between static and dynamic efficiency. Thus, discussions of this trade-off depend on a robust understanding of markups.

The link between returns to scale and markups is not merely an artifact of endogenous growth models. Returns to scale and markups are determined simultaneously, and under free entry with no sunk costs, symmetric firms will equate their markup ratios to their returns to scale in equilibrium.[5] With heterogeneous

[1] Arrow and Debreu [1954]
[2] Arrow [1951]
[3] For canonical examples see Romer [1990] and Aghion and Howitt [1992]
[4] In the Schumpeterian framework: Aghion et al. [2019], De Ridder [2024]. In the Endogenous technological change framework: Boucekkine et al. [2017], Latzer et al. [2019], Etro [2020]
[5] See Zhelobodko et al. [2012] for a fairly general example of this phenomenon. Intuition with respect to the generality can be

firms à la Melitz an equivalent relation holds with the "marginal" firm, an important mechanism in canonical trade models. Sunk entry costs or entry restrictions of any type push the markup ratio above the degree of returns to scale. With constant returns to scale and no barriers to entry, firms would continue to enter until markups are fully eroded, and one is left with the perfect competition benchmark of a zero markup (and a markup ratio of one).

In the short run, markups and imperfect competition also play a pivotal role in sticky price models. Indeed, sticky prices require price-setting firms, and only firms with market power are price setters. Additionally, strategic interactions that can prevent firms from adjusting prices directly to the long run equilibrium—slowing down *aggregate* price adjustment—depend on market power.[6]

An important prediction of sticky price models in the absence of sticky wages (or if key parameters make sticky wages a minor feature of a model) is that of countercyclical markups. Because of this, empirical investigations of the cyclicality of markups have a history nearly as long as that of sticky price models[7]. Despite this long history, there is little consensus on the truth of the matter, with examples of evidence supporting countercyclical markups[8] as well evidence to the contrary.[9] Much of this disagreement seems to come down to how movements of markups are inferred. Historically, the need to infer rather than observe firm level markups represents the most severe challenge to this literature.[10] Variation in markups is often inferred from proxies[11], sector aggregates[12], or theoretically correlated variables. As examples of theoretically correlated variables, Rotemberg and Woodford [1999] look at joint movements in the ratio of sales to inventories and the discounted growth rate of output prices, while Hall [2012] leverages the link between markups and advertising.

The connection between markups and advertising is a fundamental one, as has been understood by economists as early as Chamberlain [1933][13], who used the existence of advertising as a key piece of evidence for the need to explicitly model market power. The key to this connection is that advertising allows firms to increase sales *holding prices constant*. Then, since firms must still produce additional goods if they sell more,

seen by noting the markup ratio is given by $\mu \equiv \frac{P}{MC}$, while degree of returns to scale are given by $\gamma \equiv \frac{\partial log Y(X)}{\partial log X} = \frac{Y'X}{Y} = \frac{P_X X}{Y}\frac{Y'}{P_X} = \frac{AVC}{MC}$. Then with free entry profits equal 0, ie $AVC = P$, thus $\gamma = \frac{AVC}{MC} = \frac{P}{MC} = \mu$. Where Y denotes output, X inputs, P_X the price of inputs, P the price of output, AVC average variable cost, and MC marginal cost.

[6] See Kimball [1995] for a further discussion of the importance of imperfect competition and returns to scale in this class of models. For details on *real rigidities*, the strategic interactions that prevent prices from jumping directly to their long-run equilibrium see Ball and Romer [1990].

[7] Gordon [1981] is the earliest example of a sticky price macroeconomic model that this author is aware of is, while Domowitz et al. [1986], published 5 years later is the earliest example of an investigation of the cyclicality of markups that this author is aware of.

[8] Ex: Galeotti and Schiantarelli [1998], Rotemberg and Woodford [1999], and Bils and Kahn [2000].

[9] Ex: Domowitz et al. [1986], Hall [2012] and Nekarada and Ramey [2020].

[10] See Nekarada and Ramey [2020] for a more in-depth discussion of the challenge of measuring markups in this context, as well as a more thorough history of thought and literature review.

[11] Ex: labor share, as used by Nekarada and Ramey [2020].

[12] Ex: Galeotti and Schiantarelli [1998] and Domowitz et al. [1986]

[13] See Bagwell [2007] for a thorough history of thought regarding the economics of advertising and marketing.

the benefit from an additional sale generated by additional advertising is given by the difference between price and marginal cost: the markup. Chamberlain's observation was then that under perfect competition there is no incentive to advertise. That is, when marginal costs are equal to prices there are no profits to be gained by increasing sales, and if firms are price-takers, there is no gain on the price front either. More generally, the optimal level of advertising expenditure is proportional to the markup[14], thus one cannot think critically about advertising without thinking about markups and market power.

Finally, over the past decade, macro-economists have taken special interest in the evolution of competition at an aggregate level, claiming there has been a "decline of business dynamism", by which they mean a rise in market concentration, a rise the profit share, a decline in labor's share of income[15] and preeminently, a rise in aggregate markups.

1.2 Markup Estimation Techniques

Markups are central to many topics and questions economists care deeply about. Given their importance, macroeconomics and industrial organization (IO) have long traditions of markup measurement.

In IO, the rise of structural estimation methods under what has been dubbed the new empirical industrial organization (NEIO), has endowed researchers with a toolkit to measure markups. A major aspect of the NEIO was the refocusing of IO toward intra-industries studies. (See Bresnahan [1989].) As a result, following the NEIO, markups in IO are primarily estimated within sectors, at the product level. In these contexts, economists often have much better data on the demand side than they do on the supply side. In response, many tools are designed to measure the price elasticity of demand—which governs the optimal markup in standard models. The price elasticity of demand is measured at the product level. A key example is Rosse [1970] where markups are estimated entirely from the demand side, without any supply-side data. As such, the most powerful tools of markup estimation in the NEIO toolkit is the "demand estimation" approaches of Berry et al. [1995] (henceforth BLP) and Nevo [2000] or modification of their approach. As the name suggests, the "demand estimation" (or BLP approach as it is often called), relies on estimating demand parameters in a discrete choice framework. Because of the discrete choice framework of BLP, it is inappropriate for products for which consumers make a more continuous quantity choice. Thus, cannot be used for the full range of products, meaning in turn that one cannot obtain aggregate or broad-sector markup estimates using the BLP approach. These limitations have not fully closed the door to macro-esq investigations using demand estimation techniques. Motivated by the discussion of the decline of business dynamism, Döpper et al. [2021]

[14] See Dorfman and Steiner [1954]. This is the result leveraged by Hall [2012].

[15] On the rise of market concentration: Autor et al. [2017a] Autor et al. [2017b]; On the rise of the profit share of GDP Gutiérrez and Philippon [2016]; On the declining labor share: Karabarbounis and Neiman [2013], Elsby et al. [2013], Lawrence [2015]

departs from the usual intra-industry focus and uses a demand estimation approach to estimate markups using a plausibly representative sample of consumer products, documenting a substantial fall in the average price elasticity of demand, and correspondingly a rise in the optimal markup. This tends to back up results based on firm-level markups.

In macroeconomics, Hall [1988] was among the first to provide clear evidence that the perfect competition benchmark does not give an accurate picture of the US economy at the macro level. Before that, the perceived difficulty of measuring markups led most macroeconomists to stick with the perfect competition benchmark. Following Hall [1988] modern results on markups at a sectoral or aggregate level tend to rely on the "Production function approach" to markup estimation.[16] This research agenda, spearheaded by De Loecker et al. [2020] (Henceforth DEU) documents a substantial rise in market power between the 1980s and 2016.

DEU builds off tools of Hall [1988] to circumvent the limitations of the BLP approach, developing what has been dubbed the "production function" (PF) approach. PF estimation begins by relating the markup to the output elasticity of a variable input, which is then estimated following Ackerberg et al. [2015] (henceforth ACF). In recent years the production function approach has faced much criticism. While the issues will be discussed in depth in subsequent sections, the major criticisms are related directly to production function estimation. A key criticism, raised early on by Klette and Griliches [1996], is that production function estimation should in principle utilize quantity data, but due to data limitations is typically estimated using revenue (price $\times$ quantity) data as if it were quantity data.[17] Another issue is unobserved input utilization[18] which further threatens identification of the production function parameters. As is known in the literature, the direction and magnitude of these biases are unclear, and they cannot be properly controlled or instrumented away with the available data.[19] *A priori*, there is no reason to believe that the sum of these biases' is constant over time, making it unclear what conclusions can be drawn regarding the trend in markups.

In addition to the NEIO-BLP tradition, and Hall-DEU tradition, there is a third approach to markup estimation: the cost accounting (CA) approach. Despite being seldom used in recent years, the cost accounting method has the oldest history, with use in macro dating back a least to Domowitz et al. [1986], with a much larger tradition in the IO literature as one of the major tools of the Structure-Conduct-Performance paradigm (SCPP)[20]. The cost accounting approach fell out of favor in industrial organization with the NEIO in favor of approaches better suited to the data demands and the field's goals. It is important to note that

[16] See Hall [1988], Loecker and Warzynski [2012], De Loecker and Scott [2016], De Loecker et al. [2020]
[17] See also Bond et al. [2021] and Ridder et al. [2022]
[18] See Basu [1996], Basu and Kimball [1997], and Basu et al. [2006]
[19] In the case of the error induced by using revenue data, Bond et al. [2021] argues that no valid instrument can exist. Since demand is a function of price, any informative instrument of quantity must also shift output prices, violating the exclusion restriction.
[20] See Bresnahan [1989] for an overview of the history of thought for the Structure-Conduct-Performance paradigm.

this transition was motivated almost purely by pragmatic considerations. Though there have been some theoretical criticisms[21], this author is not aware of any prior empirical tests of the cost accounting approach to markup estimation or of any systematic comparison to alternative methods.

The cost accounting approach has one key assumption: constant marginal costs, which implies marginal cost equal to average variable costs, allowing the markup ratio to be estimated as $\mu \equiv \frac{P}{MC} = \frac{Revenue}{Total\ Variable\ Cost}$. While that assumption may be only an approximation, the advantages to this approach are apparent: it is highly transparent, making any potential pitfalls clear and pointing to relevant dimensions for robustness tests. Moreover, the cost accounting approach is simple and data-thrifty allowing for implementation in many cases where alternative methods are infeasible.

This paper argues that the lack of recent attention to the cost accounting approach (CA) is unwarranted. First, given the value of evidence regarding markup and more broadly market power, researchers should pursue all possible angles in pursuit of a robust and clear understanding. Second, the data-thrifty nature of CA means it can be implemented in situations where the others cannot, and in contexts where PF and CA are both feasible, CA seems to perform better. Furthermore, none of the three methods are without faults, so ideally one should compare results from many methods. As CA is the easiest to implement, it should certainly be on the list of methods compared.

1.3 Organization

The organization of this paper is as follows: Section 2 introduces the main data set and outlines the cost accounting (CA) and production function (PF) approaches to markup estimation. Section 3 compares the predictions of the two approaches. Section 4 discusses the pitfalls and mechanical similarities between the two measures. Section 5 formalizes the connection between advertising and markups and develops a test based on the Dorfman and Steiner [1954] advertising equation. Section 6 considers several applications that take advantage of the strengths of the cost accounting approach. Finally, section 7 concludes with a brief discussion.

2 Markup Estimation

2.1 Data

There are two considerations for data selection: coverage and being able to compare results to the literature. With these considerations in mind, COMPUSTAT Fundamentals Annual is the clear choice, offering annual

[21] See Bresnahan [1989] for a discussion of these criticisms

accounting data for all publicly traded firms in the United States since 1950. This data set has been widely used in the macroeconomics, accounting, and finance literatures. It is also the most widely used data set in the macroeconomic markup estimation literature.

Although COMPUSTAT is not a direct analog for the universe of firms in the US, it is the closest feasible alternative. It offers coverage across all two-digits NAICS industry codes, covering around thirty percent of employment[22] and forty percent of revenue[23] in the United States. Relative to the universe of firms, publicly traded firms are disproportionately older, larger, and employ more capital.

COMPUSTAT offers hundreds of firm-level accounting variables. Implementation of the cost accounting and production function methods requires data on revenue, input costs, and capital stock, and later (for validation) advertisement spending. Total revenue is reported as Net Sales/Turnover (SALE). Generally Accepted Accounting Principles (GAAP) divide total expenses (XOPR) into two broad categories, Cost of Goods Sold (COGS) and Selling, General and Administrative Costs (SGA). It is standard in the literature to take SGA as a measure of total fixed costs, and COGS as a measure of variable costs. The capital stock is taken as the level of reported gross property plants and equipment (PPEGT), as is standard in the literature.

The sample used in the next section is limited to all firms that report nonnegative values for COGS, XSGA, PPEGT, and SALES. Furthermore, the sample is trimmed by 1% at the top and bottom of the ratio of COGS to SALES and the ratio of XSGA to SALES. Summary Statistics are provided in appendix E.1.

2.2 Cost Accounting Approach

The primary hurdle to observing markups in economic data is measuring marginal costs. All three methods of markup estimation must face this challenge; each does so in a different way. The cost accounting (CA) approach assumes constant marginal cost, which makes the markup ratio equal to the ratio of total revenue to total variable cost. Beyond this one key assumption, the cost accounting needs no other big assumptions, beyond the assumption all methods share that firms are providing meaningful data. In particular, the CA approach does not rely on any other assumptions about firms' production functions. Consider an arbitrary production function $Y = Y(X)$, where X is a vector of inputs. Denoting the vector of input prices by P^X,

[22] Davis et al. [2006]
[23] Asker et al. [2015]

cost minimization implies a cost function C given as[24]:

$$C(\bar{Y}) \equiv \min_{X \in \left\{Z|Y(Z)=\bar{Y}\right\}} P^X \cdot X$$

By letting $F_c = C(0)$, without loss of generality we can decompose the cost function as:

$$C(Y) = F_c + V_c(Y)$$

Then variable cost V_c is such that $V_c(0) = 0$, and F_c denotes any and all fixed cost. Average variable cost is given by $AVC = \frac{V(Y)}{Y}$, and marginal cost is given by $MC = \frac{\partial C(Y)}{\partial Y} = \frac{\partial V(Y)}{\partial Y}$. Then if we assume a constant marginal cost $V(Y) = MC \times Y$ we have $MC = AVC$. Under this assumption, marginal costs are observed so long as variable costs are observed.

To match the data total variable costs are taken as $V_c = COGS + PPEGT \times r$, where COGS and PPEGT come directly from accounting data, and r denotes the user cost of capital. User cost of capital is attained following the standard procedure from De Loecker et al. [2020]. That is, user cost of capital is taken as $r_t = I_t - \Pi_t + \triangle$, where I_t denotes the interest rate taken to be federal funds rate, Π_t denotes the inflation rate, taken from CPI, and $\triangle$ is the sum of depreciation and risk premium, taken as 12%. Then the cost accounting markup ratio is given as[25]:

$$\tilde{\mu}_{CA} \equiv \frac{PY}{V} = \frac{SALE}{COGS + PPEGT \times r} \tag{1}$$

2.3 Production Function Approach

The production function approach[26], infers marginal cost from the firms cost minimization problem for a variable input. The key insight is rooted in cost minimization. Under this condition, the markup is equal to the product of the output elasticity of a given input and the revenue share of that input. This insight is owed to Hall and allows a markup estimation from a function of the output elasticity of a variable input (production function component) and the ratio of revenue to variable cost expenditure (the accounting component).

Consider a firm that produces a quantity of output Y given a variable input V and a vector of technology

[24] Strictly speaking the existence of a cost function does not require cost minimization. The existence of any injective mapping from desired output to inputs will suffice and does not need to be observed by the economist. For example ,consider a case where when selecting inputs producers misperceive input costs as $P^{\tilde{X}}$ and act inline with some spurious constraint $F(X) = 0$, then the cost function is given as $C(\bar{Y}) = P^X \cdot \left(arg \min_{X \in \{Z'\}} P^{\tilde{X}} \cdot X \right)$, where $Z' = \left\{Z|Y(Z) = \bar{Y}\right\} \cap \{Z|F(Z) = 0\}$

[25] Throughout I will refer to the markup ratio ($\frac{P}{MC}$) as μ, and the markup ($\frac{P-MC}{MC}$) as $\mu - 1$. The Lerner index $\left(\frac{P-MC}{P}\right)$ can then be denoted as $1 - \frac{1}{\mu}$.

[26] Hall [1988], the version implemented follows from Loecker and Warzynski [2012], De Loecker and Scott [2016], and DEU

and all other inputs X^{-V} from the production function:

$$Y = F(V, X^{-V})$$

The condition relating the markup to the output elasticity and accounting component can be seen directly from the firm's first-order condition with respect to the variable input[27]. However, deriving the condition explicitly offers additional intuition about the role of its underlying assumptions. First ,let us consider the production function component. The *output elasticity* with respect to the variable input V can be defined as:

$$\theta_v \equiv \frac{\partial log\left(Y\right)}{\partial log\left(V\right)} = \frac{\partial log\left(F(V, X^{-V})\right)}{\partial log\left(V\right)} \tag{2}$$

The output elasticity θ_v then tells us the % increase in output that results from a 1% increase in the quantity of the variable input V. It is then useful to relate θ_v (which will be estimated later) to a marginal impact on revenue. It will prove useful to define the *revenue response* of variable input expenditures while holding input and output prices constant:

$$u_v \equiv \frac{P}{P_v}\frac{\partial Y}{\partial V} = \theta_v \frac{PY}{P^v V} \tag{3}$$

The revenue response u_v then tells us the \$ increase in revenue that results from a \$1 increase in expenditures on the variable input V, holding input prices (P_v) and output prices (P) constant. Note then that u_v is the product of the production function and accounting component, and will serve as the production function estimate of the markup ratio. u_v's relationship to the markup can then be understood by reinterpreting the markup ratio. The markup ratio $\mu \equiv \frac{P}{MC} = P\frac{\partial Y}{\partial C(Y)}$ can be interpreted as the increase in revenue that results from a \$1 increase in expenditures, holding prices constant.[28] That is, the only difference between the u_v and μ, is that u_v is the revenue response to expenditures on the variable input V, while μ is that via whatever channel the firm chooses to spend on at the margin.

To see the final result we must then consider the fact that under optimal price minimization, firms equate marginal returns across all variable inputs, thus $u_v = \mu$,[29] and we have the production function approach markup ratio as:

$$\tilde{\mu}_{PF} \equiv \theta_V \frac{PY}{VP^V} \tag{4}$$

[27] See Hall [1988], Loecker and Warzynski [2012] or De Loecker et al. [2020]. Additionally, a derivation is given in appendix A.
[28] Note that this assumes that firms take input prices as given in the cost function, that is firms have no monopsony power.
[29] For further exposition consider the case where firms fail to minimize costs. Let μ^* denote the optimal markup that arises from cost minimization and consider a firm that selects inputs via some sub-optimal method that yields a cost function $C(Y)$. Then holding quantity fixed we have $\mu \leq \mu^*$ and $u_v \leq \mu^*$, that is that firms are not maximizing their revenue response.

It is then clear that (4) relies primarily on an assumption of optimal cost minimization, which implies equating revenue responses across variable inputs. The revenue response of V is in turn the product of the output elasticity of V, and the ratio of revenue to total expenditures on V. The inverse of the latter $\left(\frac{P^V V}{PY}\right)$ can be interpreted as the variable input's share of revenue.

In practice, cost of goods sold (COGS) serves as the primary measure for variable input expenditures. The accounting component of the markup ratio is then calculated as the ratio of SALE (revenue) to COGS:

$$\frac{PY}{VP^V} = \frac{SALE}{COGS}$$

The output elasticity or production function component θ_v is not observed directly and will have to be estimated. As we will see, the θ_v is the primary point of discussion of the production function approach, both in terms of implementation and of its criticisms. The accounting component and its influence get much less attention but will be a primary focus of section 4.

2.3.1 Production Function Estimation

To estimate θ_v, econometricians typically follow a modified version[30] of the ACF two-step production function estimation method. I follow the version of this method developed in Loecker and Warzynski [2012] and extended in De Loecker and Scott [2016] and De Loecker et al. [2020] to allow for first and second-stage controls.

The first step involves choosing the functional form of the production function and determining the level at which parameters will be estimated (e.g., at the sector level). In this paper, I use the main approach from DEU, estimating the parameters at the two-digit NAICS sector level using a 5-year rolling window and a Cobb-Douglas production function structure.[31] That is, within a given sector, and 5-year rolling window log output is modeled as:

$$y_{it} = \theta_{v,s,t} v_{it} + \theta_{k,s,t} k_{it} + \omega_{it} + \epsilon_{it} + g_{it} \tag{5}$$

Where y denotes the log of SALES (revenue), v is the log of COGS, and k is the log of PPEGT (capital). Additionally, ω represents firm level Hicks-neutral productivity term, and ϵ captures firm specific measurement error. Because inputs are measured in units of cost (rather than raw inputs) there is also the error term g

[30] It is worth noting that the micro-foundations of ACF rely on a value-added production function, while in practice those estimating markups typically assume a gross output production function, see Gandhi et al. [2020] for a discussion of the issue and associated bias.

[31] In addition DEU (2020) includes several alternative specifications, including a translog specification, though none of them produces meaningfully different estimates. It is worth noting that attempts to estimate the parameters more granularly (either with smaller rolling windows, or more narrow sector definitions) lead to poorly behaved estimates. Given that COMPUSTAT is the largest data set with coverage across all sectors this issue is unavoidable.

which captures unobserved variation in input prices. Following DEU, I assume that the unobserved variation in input prices is given as a function of market share z_{it}, and the product of the market share with each input, ie:

$$g_{it} = g(z_{it}, k_{it} \times z_{it}, v_{it} \times z_{it}) \tag{6}$$

The first stage serves to purge the measurement error ϵ, by estimating the non-parametric function function:

$$y_{it} = \phi(v_{it}, k_{it}, z_{it}) + \epsilon_{it} \tag{7}$$

The key assumption that allows for the construction of (7) lies in the assumption that the flexible input v is a function of ω in equilibrium, that is:

$$v_{it} = f(\omega_{it}, k_{it}, z_{it})$$

Then if we assume that f is invertible in ω, and denoting that inverse as f^{-1} we have:

$$y_{it} = \theta_{v,s,t} v_{it} + \theta_{k,s,t} k_{it} + \omega_{it} + f^{-1}(v_{it}, k_{it}, z_{it}) + \epsilon_{it} + g_{it}$$

Then defining $\phi(v_{it}, k_{it}, z_{it}) \equiv \theta_{v,s,t} v_{it} + \theta_{k,s,t} k_{it} + \omega_{it} + f^{-1}(v_{it}, k_{it}, z_{it}) + g_{it}$ yields (7), and $\hat{\phi}$ asymptotically approaches y_{it} and is assumed to be free of measurment error. The variable of interest θ_v is not identified in this first stage. θ_v is identified in the second stage with the help of an added assumption of an $AR1$ process on ω, ie $\omega_{i,t} = \rho \omega_{i,t-1} + \xi_{it}$.

To build the moment conditions of the second stage, first stage estimates are used to construct ω and ξ as functions of $\theta \equiv \{\theta_{v,s,t}, \theta_{k,s,t}\}$ and g_{it}. That is:

$$\begin{aligned}
\hat{\omega}_{it}(\theta; g) =& \hat{\phi}_{it} - \theta_{k,s,t} k_{it} - \theta_{v,s,t} v_{it} - \hat{g}_{it} \\
\xi_{it}(\theta; g) =& \hat{\omega}_{it}(\theta; g) - \rho \hat{\omega}_{i,t-1}(\theta; g)
\end{aligned}$$

g is taken to be a flexible function of the form given in 6.[32] Within the estimation routine, ρ will be estimated by projecting ω on its lag, given $\{\theta, g\}$[33]. Then estimates of θ (and the flexible function of g) are attained

[32] See De Loecker and Scott [2016] for justification of this form, and discussion of identification in the presence of second-stage controls.

[33] That is ρ is calculated at each step of the optimization routine to solve for $\{\hat{\theta}, \hat{g}\}$.

from numerically solving the moment condition:

$$\mathbb{E}\left(\xi_{it}\begin{bmatrix} v_{it-1} \\ k_{it} \\ Z_{it-1} \end{bmatrix}\right) = 0 \tag{8}$$

Where Z_{it-1} is a vector that contains the collection of polynomial terms of the set $\{z_{it-1}, k_{it-1} \times z_{it-1}, v_{it-1} \times z_{it}\}$, and is needed to ensure identification of the flexible function g.[34]

3 Comparing Markup Estimates

Despite its simplicity, this author is not aware of any prior systematic comparison of the predicted markups from the cost accounting approach with those from the production function approach. While the production function approach heavily relies on accounting data for both its production function and accounting components, the two methods diverge significantly in their assumptions and are subject to different criticisms. This raises the possibility of meaningful discrepancies between their estimates.

Figure 1 compares the two markup measures for 1,000 sample points, with an OLS trend line fitted to the full sample. The strong alignment along the 45-degree line suggests a high degree of similarity between the two measures. The plot reveals a strong correlation between the two measures, with a coefficient of 0.877. The OLS trend line's slope is effectively 1, and its intercept is near 0, indicating that, on average, the two measures yield nearly identical results. That being said, a large cluster of points lies below the 45-degree line. Visually, there is an apparent lower bound on the ratio of the two markups, which in part drives the fact that the production function markups are slightly larger on average. For a firm in sector s, denoting the production function markup estimate, and corresponding output elasticity estimate as $\mu_{PF,i,t}$, and $\hat{\theta}_{v,s,t}$[35] respectively, it is straightforward to deduce the relation between the two estimates as:

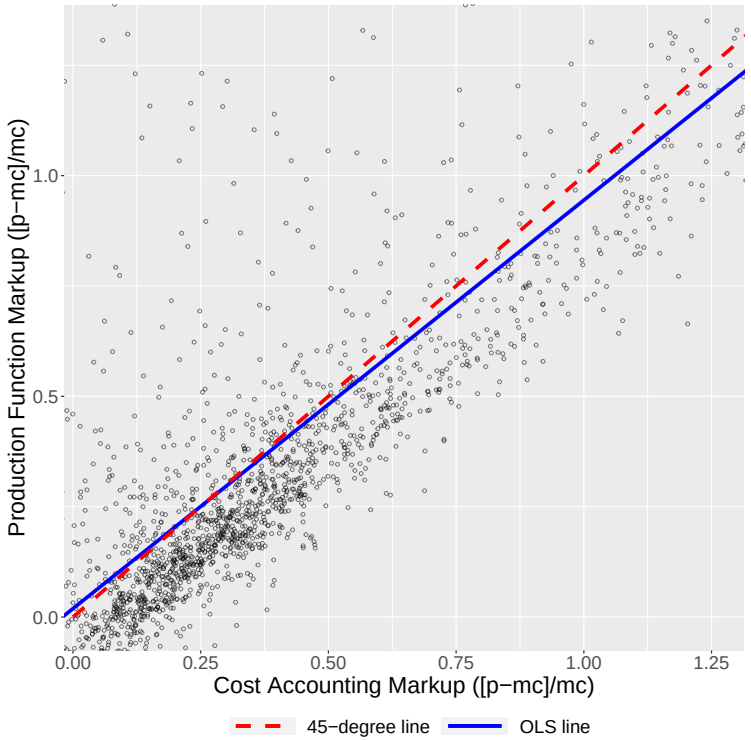
$$\frac{\tilde{\mu}_{PF,i,t}}{\tilde{\mu}_{CA,i,t}} = \hat{\theta}_{v,s,t}\left(\frac{COGS_{it} + PPEGT_{it} \times r_t}{COGS_{it}}\right) \tag{9}$$

The visual pattern seen in the plot then arises because $\frac{\tilde{\mu}_{PF}}{\tilde{\mu}_{CA}}$ is unbounded above, but bounded below. From (9) note that the ratio tends to infinity as the ratio of capital to non capital variable costs $\left(\frac{PPEGT_{it} \times r_t}{COGS_{it}}\right)$ tends to infinity. Then a firm with a relatively high level of capital expenditures will admit a higher measure under the production function approach. Additionally from (9) note that as the capital expenditure share

[34] See De Loecker and Scott [2016] for the derivation of the moment condition.

[35] Recall that $\hat{\theta}_{v,s,t}$ is estimated at the sector time level in 5-year rolling windows, the s, and t subscripts are included to make this clear.

Figure 1: Markup Comparison: Scatter Plot with 1000 Sample Points



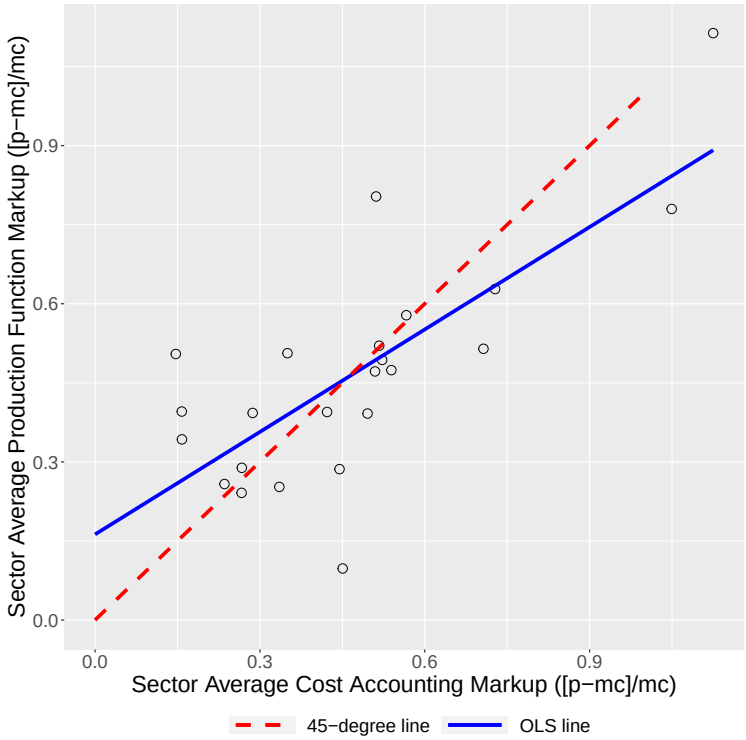
Notes: Cost accounting (CA) and production function (PF) markup estimates for a random sample of 1,000 firms. The OLS line is fitted from the full sample. Full sample correlation coefficient: 0.877.

tends to 0 we attain a lower bound, ie: $\frac{\tilde{\mu}_{PF,i,t}}{\tilde{\mu}_{CA,i,t}} \geq \hat{\theta}_{v,s,t}$. This relationship underscores the role of capital in driving differences between the measures. Note that while this lower bound technically varies at the sector time level (since output elasticities vary at this level), it does not vary much as the range of the elasticity estimates is quite tight.

Figure 2 presents the scatter plot of sector averages according to each markup measure, with sectors again defined at the two-digit NAICS level. At the sector level, the correlation coefficient drops to 0.630, and the OLS line deviates substantially from the 45-degree line. This indicates that cross-sector variation contributes significantly to the differences between the two measures. This suggests that the variation in the difference between markup measures is largely driven by variation across sectors, or equivalently that the correlation is largely driven by variation within sectors.

Figure 3 compares the average markup by year across the two markup measures. The two annual averages closely align in both level and trend, demonstrating the robustness of prior evidence regarding the rise in aggregate markups. Additionally, this alignment suggests that CA provides a viable alternative for capturing long-term markup trends. Despite its simplicity, it matches the production function approach in measuring aggregate markups, as evidenced by the close agreement in their trends and levels.

Figure 2: Markup Comparison: Scatter Plot



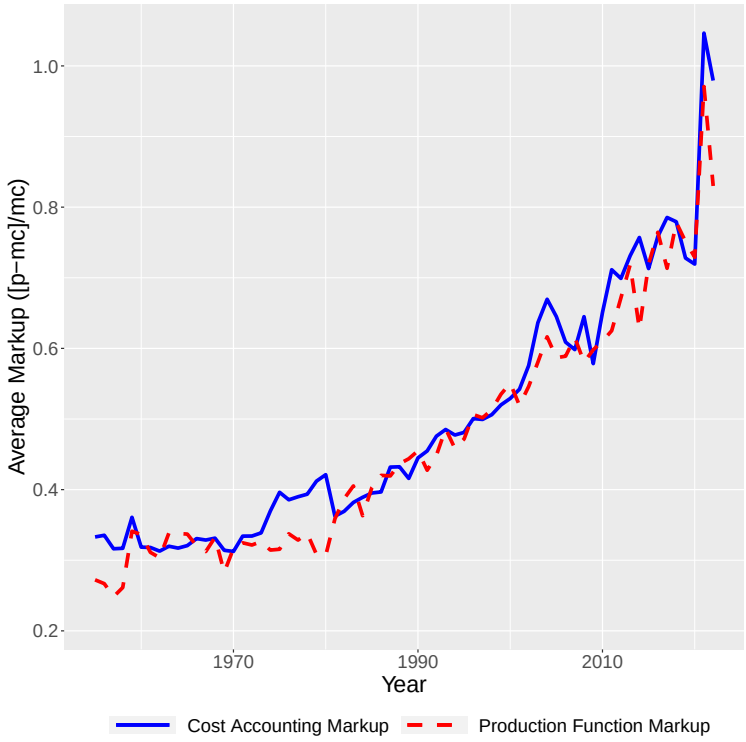
Notes: Sector averages are constructed at the two-digit NAICS level. Correlation coefficient: 0.630.

The two averages are highly correlated, with a correlation coefficient of 0.973. Sales-weighted averages and averages by sector are presented in appendix B. The results for the sales-weighted averages are similar, and they admit a correlation coefficient of 0.880. The headline result of DEU, the markups rose between 1980 and 2016 (the last year included in their analysis) is robust both qualitatively, and quantitatively robust to the cost accounting approach of markup estimation. Theoretical differences between the two measures—such as the reliance of the production function approach on an estimate of the output elasticity—make their agreement particularly notable and lend credibility to the robustness of the observed trend. Furthermore, with the additional 6 years of data, the rise in markups seems to persist for both measures. It is worth noting, that given the unclear effects of the COVID-19 recession, one may be hesitant to forecast the future of this trend, but economists and policymakers should continue to follow this as time goes on.

The second major observation from DEU was that the rise in markups was almost solely driven by a rise in "superstar firms"—large, efficient firms that exhibit large markups. This can be seen by an increase in the right tail of the markup distribution over time.

Figure 4 compares the densities in 1980, and 2016 (the years that DEU marks as the primary period for the rise in markups) as well as those in 1955 and 2022, the first and last year in the data set used in this

Figure 3: Markup Comparison: Annual Averages



Notes: Average annual markups for CA and PF from 1955 to 2022. Correlation coefficient: 0.973.

paper. Overall, there is a stark similarity between the distributions of the two markup measures for all years.

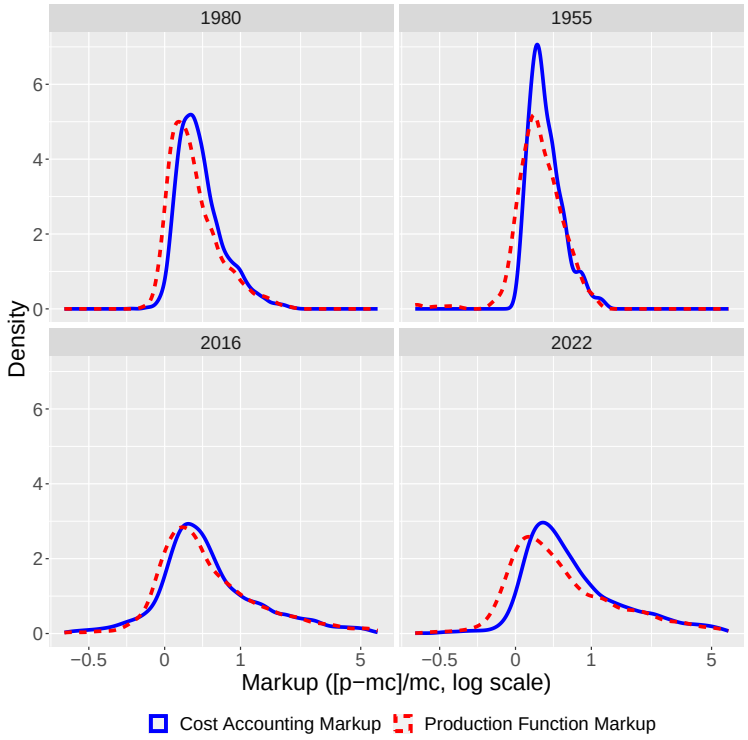
As with the production function markups, we see the distribution elongate over time. Particularly the right tail of the distribution is substantially fatter in the later years than in the earlier ones. As with the result on the rise in aggregate markups, the DEU result on the rise of superstar firms appears robust to the cost accounting approach of markup estimation. Comparison of densities across the entire sample are given in appendix B.

4 Biases and the Role of SALE/COGS

Despite the fundamentally different assumptions underpinning the cost accounting and production function approaches, the strong similarity between their estimates is surprising and warrants closer examination. This raises two key questions: To what extent is the similarity mechanical, and should it inspire confidence or skepticism in the two measures? I approach these questions first with a discussion of the issues and potential pitfalls of each estimation technique, and then through this lens home on the mechanical relation between the two.

As shown in equation (4), the production function markup estimate consists of two components. The

Figure 4: Markup Comparison: Markup Densities



Notes: This density plot displays the distributions of CA and PF markups for the years 1980, 2016, the two years discussed in DEU as well as 1955, and 2022, the first and last years in the sample. Both methods show a notable elongation of the right tail over time, indicating an increase in the number of firms with high markups. The similarity in these distributions suggests that the rise of "superstar firms" with high markups is robust to the choice of markup estimation method.

first is the production function component, represented by the output elasticity θ_v. The second is the accounting component, calculated as the ratio of revenue to expenditures on the variable input $\frac{SALE}{COGS}$. The main contention with respect to cost accounting markups comes down to its key assumption: constant marginal cost as a good approximation. In contrast, the primary points of contention with respect to the production function technique relate to the production function estimation component. These differing points of contention raise the question of whether the observed similarity between the measures reflects robust estimation or shared reliance on accounting data. Equations (4) and (1) reveal that the similarity between the two measures could arise directly from their shared reliance on the accounting variables $SALE$ and $COGS$ which are not related to the main criticisms of either method.

4.1 Criticisms of Cost Accounting Markups

Bresnahan [1989] is the generally agreed upon citation for rejection of any accounting approach to markup estimation. Despite this, discussion of accounting approaches to markup measurement as such is somewhat

lacking[36]. Chapter 17 of the Handbook of Industrial Organization, is largely dedicated to outlining the new empirical industrial organization (NEIO) as a replacement to the "structure-conduct-performance paradigm (SCPP)", and discussion of accounting markups is limited to the following:

> *The NEIO is partly motivated by dissatisfactions over three maintained hypotheses in the SCPP: (i) economic price-cost margins (performance) could be directly observed in accounting data ... The NEIO is an attempt to continue the use of such evidence while returning to the study of single (or related) industries.*

The second portion of this quotation is included to offer further context to the motivation of the NEIO. It is worth noting that the question of the validity of accounting markups in no way undermines the value of NEIO's toolkit of markup estimation since the measurement of accounting markups does not provide a way to conduct counterfactual experiments without additional structure and data. Furthermore, demand estimation techniques can be implemented with demand side data alone, and are the only feasible method when supply side data is unavailable[37]. Lastly, the emphasis on product-level analysis in NEIO rules out the cost accounting approach and the production function approach in most contexts.

The assertion that marginal costs (and thus markups) cannot be observed in the data is not without merit. The primary hurdle is in inferring marginal costs from variable costs. The crux of this issue is that if the constant marginal cost assumption fails, making this assumption introduces a multiplicative error equal to the degree of returns to scale of the production function when ignoring fixed costs:

$$\frac{AVC}{MC} = \frac{V}{Y \times MC} = \frac{V}{Y}\frac{\partial Y}{\partial V} = \frac{\partial log(Y)}{\partial log(V)} \equiv \Gamma, \tag{10}$$

where Γ can be thought of as the degree of returns to scale of the variable cost function. Trivially, when the variable cost function is linear we have $\Gamma = 1$ and observe marginal costs directly. The cost accounting approach implemented here assumes constant marginal costs and $\Gamma = 1$. It is important to note that this is not an assumption of constant returns to scale, as the inclusion of fixed costs implies increasing returns to scale. There is also the secondary consideration of properly observing variable cost, this is discussed in appendix C.2.

4.2 Criticisms of Production Function Markups

The production function approach has faced increasing criticism in recent years, with a particular focus on issues related to unobserved prices, variable input utilization, and sector-level heterogeneity. These two

[36] DEU's discussion of such an approach is limited to referencing this work.
[37] See Rosse [1970] where markups are estimated from demand side data alone

criticisms predate DEU, emerging in response to Hall [1988]. These criticisms highlight several potential sources of bias, which could lead to either overestimation or underestimation of the true markup, depending on the specific data and assumptions. Ex-ante there is also no reason to believe that the sum of these biases' is constant over time, calling into question the validity of any markup trend recovered from the production function approach.

The markup estimation literature is well aware of the issue of unobserved prices, first noted by Klette and Griliches [1996], with the relevance to the production function approach noted by Bond et al. [2021]. This issue arises because output estimation ideally requires quantity data, yet in practice, researchers often rely on revenue (monetary terms) as a proxy. This substitution conflates changes in output with changes in prices, potentially biasing the estimated output elasticity. When ignoring this issue, what is estimated is in principle closer to a revenue elasticity rather than an output elasticity, from the simple fact that the left-hand side variable is revenue rather than output. Even in the simplest case, where the only unobserved variable is output prices, Ridder et al. [2022] shows that the production function approach yields a markup that could be biased in either direction, though in practice it seems to induce a downward bias.

The markup literature has taken less note of the unobserved variable input utilization. Basu [1996], Basu and Kimball [1997], and Basu et al. [2006] point out that unobserved variable utilization causes an upward bias in Hall's estimate of output elasticity. The problem of unobserved utilization arises when firms adjust their effective use of inputs (e.g., running machinery for longer hours) without corresponding changes in reported input quantities. If input utilization correlates with other variables, this can lead to an overestimation of the output elasticity, as output variation due to utilization is incorrectly credited to input variation.

The criticism brought up by Foster et al. [2022] also deserves discussion. They find that production function markup estimates are sensitive to the definition of sectors, ie the level at which output elasticities are estimated. This suggests that output elasticities vary within the two-digit NAICS sectors—which makes estimates for the broader two-digit sectors vulnerable to reallocation of relative production between subsectors that have different production function parameters. These three issues will be explored in greater depth in the next subsection, and additional concerns are discussed in appendix C.1.

4.2.1 Biases in The production function component

Let me analyze three major challenges that threaten the identification of output elasticity: unobserved prices, unobserved variable input utilization, and variation in production technologies within sectors. The issue with unobserved prices is twofold: rather than observing output Y, the econometrician observes revenue $R \equiv PY$, and rather than observing the variable input V, variable input expenditures are observed $\tilde{V} \equiv P^V V$. Next, issues of unobserved utilization arise when firms have an additional lever on production where they can

increase production without increasing the inputs observed by the econometrician. Then the production function takes the form $Y = F(U^v V; X^{-V})$ where U^v denotes the effective rate of utilization [38] Finally, technological heterogeneity is perhaps most salient in the Cobb-Douglas case, where it can be modeled by allowing θ_v to vary within sector, denoting firm-level measures by $\theta_{v,i}$

To simplify the analysis, I assume no measurement error and independent, identically distributed (IID) productivity shocks[39]. Furthermore to emphasize the role of the variable input consider a production function with only one input. Then denoting log revenue, log input prices, log variable expenditures, log input prices, and log utilization by $r, p, \tilde{v}, p^v$, and u^v respectively we have:

$$r_i \equiv y_i + p_i = \theta_{V,i}\left(\tilde{v}_i + u_i^v - p_i^v\right) + \omega_i + p_i \tag{11}$$

With these simplifications the moment condition analogous to (8) is:

$$\mathbb{E}\left(\left[r_i - \hat{\theta}_v \tilde{v}_i\right]\tilde{v}_i'\right) = 0 \tag{12}$$

Then from (11) we have $r_i - \hat{\theta}_v\tilde{v}_i = \left(\theta_{v,i} - \hat{\theta}_v\right)\tilde{v}_i + \theta_{v,i}\left(u_i^v - p_i^v\right) + p_i$. Then estimating (12) when the true DGP is given by (11) is asymptotically biased. This can be seen directly by combining (11) and (12) to see that

$$\hat{\theta}_v = \frac{\mathbb{E}\{\theta_{v,i}\tilde{v}_i\tilde{v}_i'\}}{\mathbb{E}\{\tilde{v}_i\tilde{v}_i'\}} + \frac{\mathbb{E}\{u_i^v\tilde{v}_i\tilde{v}_i'\}}{\mathbb{E}\{\tilde{v}_i\tilde{v}_i'\}} + \frac{\mathbb{E}\{[p_i - p_i^v]\,\tilde{v}_i\tilde{v}_i'\}}{\mathbb{E}\{\tilde{v}_i\tilde{v}_i'\}} \tag{13}$$

$$\frac{\hat{\theta}_v}{\mathbb{E}\{\theta_{v,i}\}} = 1 + \frac{\mathbb{E}\{[\theta_{v,i} - \mathbb{E}\{\theta_{v,i}\}]\,\tilde{v}_i\tilde{v}_i'\}}{\mathbb{E}\{\theta_{v,i}\}\mathbb{E}\{\tilde{v}_i\tilde{v}_i'\}} + \frac{\mathbb{E}\{u_i^v\tilde{v}_i\tilde{v}_i'\}}{\mathbb{E}\{\theta_{v,i}\}\mathbb{E}\{\tilde{v}_i\tilde{v}_i'\}} + \frac{\mathbb{E}\{[p_i - p_i^v]\,\tilde{v}_i\tilde{v}_i'\}}{\mathbb{E}\{\theta_{v,i}\}\mathbb{E}\{\tilde{v}_i\tilde{v}_i'\}} \tag{14}$$

From (13) $\hat{\theta}_v$ converges to a weighted average level $\theta_{v,i}$ plus two bias terms, arising from unobserved variable utilization (and input prices) and from unobserved output prices respectively. Dividing by $\mathbb{E}\{\theta_{v,i}\}$ reveals the bias of $\theta_{v,i}$ relative to the average output elasticity. The terms on the right-hand side of equation (14) then quantify this multiplicative bias on $\hat{\theta}_v$. Equation (14) highlights that correlations between the output elasticity and the product of current and lagged variable input expenditures introduce an additional source of bias. Given Cobb-Douglas production functions, the expenditure share of a given input is proportional to its output elasticity. This then makes for a positive correlation between $\theta_{v,i}$ and both $\tilde{v}$ and its lag. In turn, the first term on the RHS of (14) is positive. Next, a positive correlation between utilization and variable input expenditure, together with persistent demand, makes the second term positive. Finally, as noted in Ridder

[38] Allowing for unobserved input utilization also has implications for the Hall markup rule (4) and for the cost accounting markups. This is discussed in appendix.C.1.1 and C.2.1 respectively.

[39] Alternatively, we can assume that the first stage perfectly purges all measurement error and that TFP follows an AR(1) process with known parameters.

et al. [2022], the last term cannot be unambiguously signed, but given a downward-sloping demand function and persistent demand shocks, the most reasonable signing for that final term is negative, as their simulation and empirical exercises suggest.

4.3 The Role of the Revenue to Non-Capital Variable Costs Ratio SALE/COGS

The mechanical connection between production function and cost accounting markups arises entirely from the shared accounting component, $\frac{SALE}{COGS}$. From equations (1) and (4) the (log) production function and cost accounting markup can be decomposed as:

$$\begin{aligned} Ln\left(\tilde{\mu}_{PF,i,t}\right) &= Ln\left(\frac{SALE_{it}}{COGS_{it}}\right) + Ln\left(\hat{\theta}_{v,s,t}\right) \\ Ln\left(\tilde{\mu}_{CA,i,t}\right) &= Ln\left(\frac{SALE_{it}}{COGS_{it}}\right) + Ln\left(\frac{COGS_{it}}{COGS_i + PPEGT_{it} \times r_t}\right) \end{aligned}$$

Decomposing the markup measures into $\frac{SALE}{COGS}$ and their remaining components is valuable for two reasons: *i*) it isolates the components most susceptible to criticism, and *ii*) it highlights the central role of $\frac{SALE}{COGS}$ in driving both estimates. For PF, the questionable component is the estimate of the output elasticity. For CA, the questionable component is the "non-capital share in variable costs" $\frac{COGS}{COGS+PPEGT\times r}$, which requires an estimate of the cost of capital r. Fortunately, the issues raised by a possibly flawed measure of the cost of capital r are transparent: too high an estimate of r lowers the markup measure and makes it too sensitive to PPEGT; too low an estimate of r raises the markup measure and makes it too insensitive to PPEGT. $\frac{SALE}{COGS}$, by contrast, is as solid as accounting data get; the main issue for SALE/COGS would be possible misreporting by firms, which is an unavoidable problem with available data. The decomposition it highlights the potential importance of the common $\left(\frac{SALE}{COGS}\right)$ component. This raises the question of how much of the variation in each measure is driven by variation in $\left(\frac{SALE}{COGS}\right)$. *A priori*, it is possible that both markup measures are primarily linear transformations of $\left(\frac{SALE}{COGS}\right)$ despite their differing methodological underpinnings. Such a fact would go a long way toward explaining the similarity between the two measures. A straightforward way to facilitate this investigation is to regress the logarithm of each markup measure on $ln\left(\frac{SALE}{COGS}\right)$: $ln\left(\frac{SALE}{COGS}\right)$ explains 97.3% of the variance of $\ln(\mu_{PF})$, but only 79.8% of the variance of $\ln(\mu_{CA})$. That means that the output elasticity has an incremental R^2 of only 2.7% in explaining $\ln(\mu_{PF})$, while the non-capital share in variable costs has an incremental R^2 of 20.2% in explaining $\ln(\mu_{CA})$. In other words, the PF markup measure is, statistically, almost all SALE/COGS, while the CA markup measure also includes the non-capital share of variable costs in a theoretically appropriate way. A full regression table is included in appendix D.

To return to an earlier point, the agreement between the two markup measures can largely be attributed

to their mutual dependence on $\frac{SALE}{COGS}$. This highlights the central role of accounting data in driving both estimates directly. This also sheds light on the relative performance of the cross-sector and cross-time comparisons. Recall that taking annual averages yields the highest correlation between markup measures, while taking sector-level averages yields the lowest. As noted in De Loecker et al. [2020], output elasticity estimates exhibit minimal variation over time, with most variation occurring across sectors. Correspondingly, the non-capital share of variable costs has much more variation between sectors than over time. Thus, hence, the correlation between markup measures is highest for annual averages and lowest for sectoral averages, where the components other than SALE/COGS matter more.

5 Advertising and Markups

The aggregate predictions of estimates from the cost accounting and production function approach are qualitatively indistinguishable. While this suggests that the cost accounting approach is similar in its predictive power for underlying markups, it does not directly address the broader concern that observed variation may stem from measurement error rather than true differences in market power. Nor does it tell us which method should be preferred when both are feasible. To address these concerns, this section uses the Dorfman and Steiner [1954] advertising equation, which links advertising expenditure to markups, to evaluate the signal-to-noise ratio of each measure.

The Dorfman-Steiner equation formalizes the link between advertising and markups, demonstrating that advertising expenditures provide a measurable signal of market power. It is this signal that forms the backbone of this section, and allows for a test with the ability to i) falsify the null that a given markup estimate carries no signal of underlying markups, and ii) compare the relative signal-to-noise ratios of the two measures.

There are three competing theories of advertising. Advertising as "informative"[40] is the idea that advertising arises as a solution to an information problem, where firms advertise to inform consumers of the quality and features of the goods they sell; firms then have an incentive to increased sales by advertising to the consumers whose preferences align with the goods they sell. Advertising as "complementary" is the idea that advertising actually raises the value of a good to the consumer. Then advertising can be a way to raise a good's effective quality. Finally, advertisement as "persuasive" [41], is the idea that advertising has no direct welfare-enhancing effects, increasing demand for the underlying good by convincing consumers to buy goods or services that do not necessarily align with their underlying preferences.[42] In each of these three theories,

[40] See Grossman and Shapiro [1984]
[41] Dixit and Norman [1978] and Dixit and Norman [1979]
[42] Even in the persuasive model, advertisement is not without its economic benefits, as advertisements help fund the media

there is a benefit to the firm from advertising. The relationship between advertising expenditure and the markup predicted by the Dorfman-Steiner equation then depends only on the *magnitude* of the benefit the firm gets from advertising, and not where that benefit to the firm comes from.

Formally, what matters for the Dorfman-Steiner relationship is that advertising increases demand *holding prices constant*. That is, the demand for a given product can be written as a function of not only the price P, but also the level of advertisement A, ie: $Y = Y(P, A)$[43]. Normalizing to measure A in units of dollars spent, we can write the flow profit for a single product firm with access to advertising as:

$$\pi = PY(A, P) - C(Y(A, P)) - A \tag{15}$$

Where $C(Y)$ denotes the total cost of producing Y units of good in the period of question, and within-period demand Y depends on price P and advertising expenditure A. Then, a well-known result is as follows:

Theorem 1 (Dorfman-Steiner (1954)) *Single product firms facing demand $Y(A, P)$ have the optimal advertising rule:*

$$\frac{A}{Y \times P} = \frac{\partial log(Y)}{\partial log(A)} \times \frac{P - MC}{P}$$

Where A denotes advertising expenditure, MC denotes marginal cost, and P denotes the price.

Theorem (1) can be seen directly by rearranging the first-order condition from maximizing (15) with respect to A. The Dorfman-Steiner equation links a firm's advertising expenditure to its markup, with higher markups incentivizing greater advertising. This relationship provides a unique opportunity to test whether estimated markups capture true market power.

To aid intuition, note that, after generating an additional sale through advertising, the firm still needs to produce the additional good, thus the return to advertising is dependent on the wedge between the sale price and the marginal cost: that is, the markup. This clarifies another well-known result, that under perfect competition, where the markup is equal to 0, there is no incentive to advertise, and the only solution to the advertising problem is A = 0. To see Theorem (1) directly, first note that with advertising in units of expenditure, a \$1 increase in advertising expenditure yields $\frac{\partial Y}{\partial A} = \frac{Y}{A}\frac{\partial \log(Y)}{\partial \log(A)}$ additional sales, and an additional sale yields a return equal to the price cost margin: $P - MC$ since the firm now needs to produce

which they appear in, acting as a subsidy in the case of print and Television programs and the main source of revenue for many online goods and social media.

[43] Taking all other relevant determinants of demand as given.

the additional good. Equating the marginal cost of an additional sale generated from advertising to the marginal benefit yields $\frac{A}{Y}\frac{\partial \log(A)}{\partial \log(Y)}$, which can be arranged as above.

More precise intuition can be gained by expressing theorem (1) in terms of the advertising share of revenue $\left(xad = \frac{A}{Y \times P}\right)$[44], the advertising response elasticity $\left(\varepsilon_{xad} = \frac{\partial log(Y)}{\partial log(A)}\right)$[45] and the Lerner index $\left(1 - \frac{1}{\mu} = \frac{P-MC}{P}\right)$[46] giving the fraction of additional revenue not eaten up by additional costs. With these definitions, the Dorfman-Steiner equation can be written as

$$xad = \varepsilon_{xad} \times \left(1 - \frac{1}{\mu}\right) \tag{16}$$

As a result, the advertising share of revenue provides a signal about the markup, and thus about the level of market power, though this signal is muddied by the advertising response elasticity ε_{xad}. The generality of this relationship is hard to overstate: the general form in theorem (1) relies only on the assumption that firms internalize the fact that advertising affects demand—an assumption without which we cannot rationalize advertising to begin with—and that firms behave rationally, which can be easily weakened by allowing for optimization error or measurement error. Beyond that, having *some* relationship between advertising and the markup only relies on the idea that the benefit of advertising depends on the markup. As long as firms internalize the fact that the return to additional sales is increasing in the markup there must be an equilibrium relationship between advertising and the markup. Thus the idea that the advertising share of revenue provides some signal of markups is quite general.

5.1 Advertising Data

To bring the Dorfman-Steiner framework to the data, I utilize the accounting variable "Advertising expenditures" (XAD), which captures firm-level spending on various promotional activities. XAD is defined as "the cost of advertising media (i.e., radio, television, and periodicals) and promotional expenses", and is taken in this analysis as representing the total level of advertising for the firms. Under generally accepted accounting principles (GAAP), XAD is included as a portion of XSGA, and thus there is no overlap in expenses included in COGS. While firms are not legally required to report XAD as a separate line item, around 30% do opt to report.

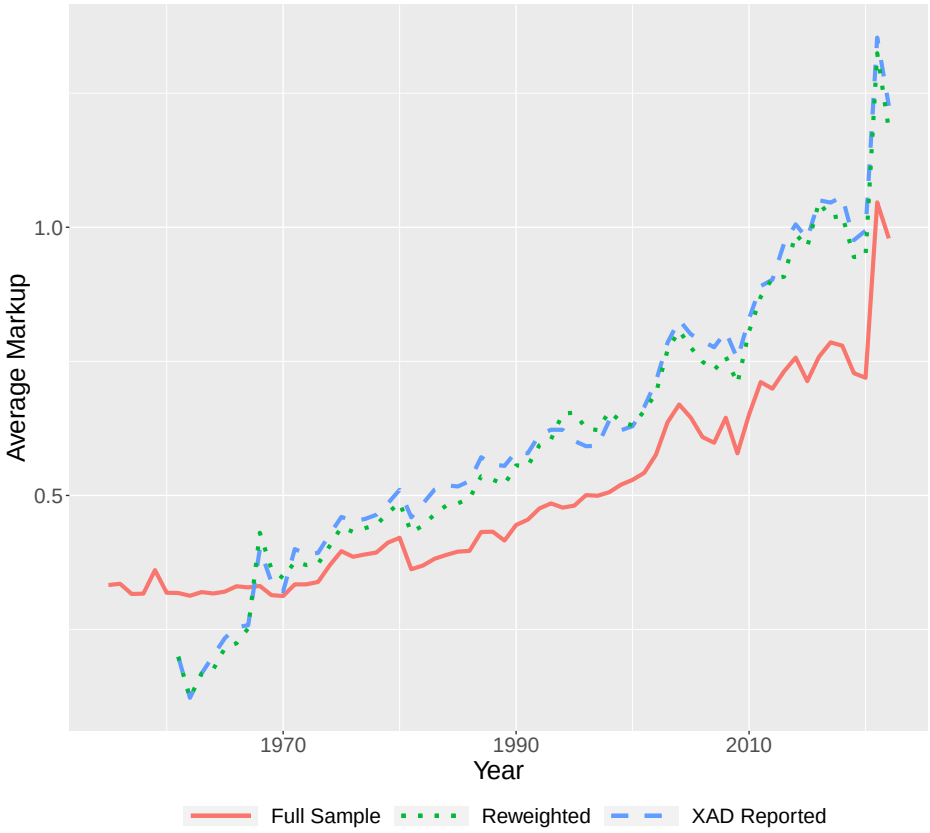
Figure 5 compares annual average markups (as measured by the cost accounting approach) across firms

[44] Which allows us to think of advertising relative to scale.

[45] A central question both for the marketing literature and for managers of firms pertains to the shape of demand while varying the quantity of advertising. This is known as the 'Advertising response function' in the marketing literature. The advertising response elasticity is then the elasticity of the Advertising response function.

[46] It is worth noting that Dorfman and Steiner take this approach one step further, noting that under monopolistic competition the Lerner index is at optimum equal to the inverse of the price elasticity of demand. It is worth noting that the underlying logic holds when prices are taken as given; thus, the above expression does not depend on optimal pricing.

Figure 5: Average Markups by Advertising Reporting



Notes: This figure plots the average cost accounting markups for firms that report advertising expenditures compared to the full sample. The re-weighted series is constructed by taking sector population weights (at the six-digit NAICS level) from the full sample and applying them to the advertising sample.

reporting advertising expenditures with the full sample. It contains the values for the full sample, the sub-sample of firms in the advertising balanced sample, and the advertising balanced sample reweighed to match the sector composition of the full sample. The reweighing is done by taking a weighted average of sector level (defined at the 6 digit NAICS level) average markups, with weights set by the number of firms in a given sector in the full sample. The XAD reported average is nearly identical to the reweighted average, indicating that with respect to markups, there is no selection by sector, thus the sample is effectively representative across sectors. Within a sector, the main selection concern is that firms engaging in higher levels of advertising might be more likely to report their advertising. If that turned out true, then (16) would imply that firms with larger markups would be more likely to report advertising. This should be borne in mind. Additional summary statistics are given in appendix E.1.

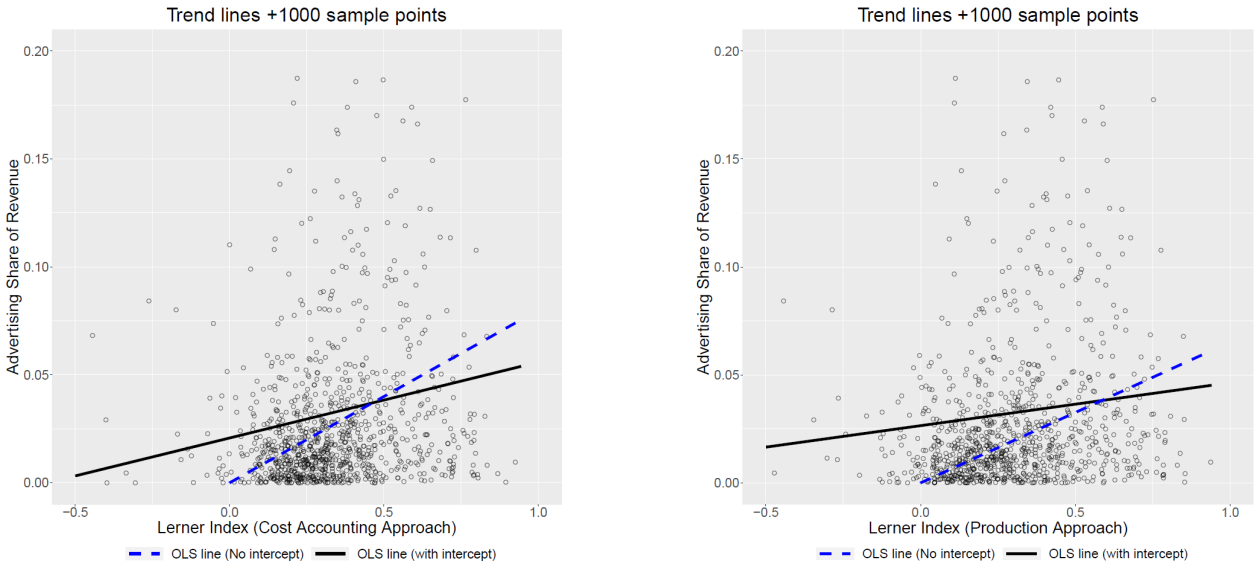
Since advertising expenditures and total revenue are observed in the data, we observe the advertising share directly. Deriving a specific test will require thinking critically about error terms, but for the moment

assume that the advertising share of revenue is observed with a mean 1, multiplicative error $e^{\epsilon_{xad}}$, then we have:

$$\tilde{xad} \equiv \frac{XAD}{SALE} = \frac{A}{Y \times P} \times e^{\epsilon_{xad}} = xad \times e^{\epsilon_{xad}} \tag{17}$$

There are various reasons we may expect an error term on $\tilde{xad}$: first, measurement error from mistakes misreporting, issues about the timing of advertising and its effects, and any gap between the accounting definition of advertising expenditures and an economic definition of advertising expenses. Then there is the possibility of optimization error in choosing advertising expenditure, which is easy to imagine in the real world. (While (16) requires optimization on the advertising expenditure margin, as noted above, it does *not* require optimal price-setting: the actual markup is what matters, not the markup a firm *should* have.)

Figure 6: Advertising Vs Markup Measures



Notes: Scatter plot of markup measures (Lerner index) on advertising share. Note that the Dorfman-Steiner equation predicts that if advertising response elasticity were constant the scatter plot would reduce to a ray from the origin.

Figure 6 presents a scatter plot of the advertising share of revenue against the Lerner index for the CA and PF markup estimates, with 1000 randomly selected observations alongside OLS trend lines with and without intercept estimated on the full sample. Both measures exhibit a positive relationship, but the stronger correlation for cost accounting markups suggests they align more closely with advertising-driven predictions of market power. The alignment suggests that the cost accounting approach is better at capturing the markups that influence firm behavior. In order to more rigorously evaluate the relationship between each markup measure and the advertising share, the next subsection develops a more formal test.

5.2 Advertising Test

While the Dorfman-Steiner equation could in principle allow for a direct test of a given markup measure using the advertising response elasticity (ε_{xad}), good firm-level estimates of ε_{xad} are unavailable[47]. Instead, this section develops a framework for comparing the relative signal-to-noise ratios of the two markup measures, with the unobserved advertising response elasticity treated as having a relationship only with the true markup and the true advertising share. This approach also makes it possible to reject the null that a given measure contains *no* meaningful signal about underlying markups.

Directly applying equation (16) to test markup measures faces two key challenges. First, revenue ($SALE$) appears in both the numerator of the markup measures and the denominator of the observed advertising share of revenue ($\tilde{xad}$). Measurement error in $SALE$ therefore introduces a mechanical negative correlation between advertising shares and markup estimates, complicating efforts to isolate the true relationship. Second, unobserved firm-level variation in the advertising response elasticity (ε_{xad}) attenuates the observed correlation between advertising shares and markups. Firms with similar markups may exhibit varying advertising shares due to differences in ε_{xad} rising from product characteristics, market conditions, or strategic differences. These two issues jointly obscure the connection between advertising behavior and markups, necessitating a robust framework to disentangle meaningful relationships from noise.

As will be shown formally below, neither challenge prevents a test of relative signal-to-noise ratios, nor does it rule out our ability to reject the null of uninformative markup estimates. Under the null hypothesis that a markup measure is uninformative[48], it should exhibit no correlation with the true advertising share. Given the negative bias introduced by measurement error in $SALE$, and attenuation bias introduced by ε_{xad}, any observed statistically significant positive correlation is a strong rejection of this null, indicating that the measure captures meaningful information about underlying markups.

Despite these challenges, the relative signal-to-noise ratios of the two measures can be benchmarked by comparing their correlations with the advertising share. This approach exploits the fact that both measures are affected similarly by the errors induced by $SALE$, and ε_{xad}. This follows from the fact that these two issues affect the relationship between the observed advertising share ($\tilde{xad}$), and the underlying markup[49], rather than the relationship between a given markup estimate and true advertising share (xad). In other words, these issues affect the correlation between $\tilde{xad}$ and cost accounting (CA) markups in the same way it affects the correlation between $\tilde{xad}$ and production function (PF) markups. As a result, neither issue obscures the *relative* ability of one measure to capture the true markup signal better than the other, and we

[47] Such data would also offer an additional way to measure markups.
[48] That is, the null that estimates $\tilde{\mu}_{CA}$ (or $\tilde{\mu}_{PF}$) are orthogonal to the true markup μ^*
[49] With measurement error on prices.

can compare relative signal-to-noise ratios.

Consider the case where markup estimates from a given method are the product of underlying markup ratios μ^*, and some source of noise. Assuming multiplicative error reflects the practical reality that noise scales with the size of the markup. This assumption simplifies the derivation of relative signal-to-noise ratios while preserving the key relationships needed for hypothesis testing. That is, I assume a mean 1, multiplicative noise term $e^{\epsilon_i^M}$. That is, for $M \in \{CA, PF\}$:

$$\log\left(\hat{\mu}_{M,i}\right) = \log\left(\mu_i^*\right) + \epsilon_i^M \tag{18}$$

Then the signal to noise ratio for a given (log) measure is defined as:[50]

$$SNR\left(\hat{\mu}_M\right) \equiv \frac{var\left(\log\left(\mu^*\right)\right)}{var\left(\epsilon_i^M\right)} \tag{19}$$

Another source of error is measurement error in the observed advertising share (as a fraction of revenue), $\tilde{xad}$. Finally, as mentioned above, there is an optimization error in the choice of advertising intensity. Overall, there are three obvious sources of error for the Dorfman-Steiner Equation: measurement errors in advertising expenditure and revenue and optimization errors. Modeling all of these errors as multiplicative errors with a mean of 1 makes the following derivations straightforward.[51] Formally, the errors are $e^{\epsilon^x}, e^{-\epsilon^s}, e^{\epsilon^a}$, where ϵ^s carries a negative sign since $SALE$ appears in the denominator of xad. Then:

$$log(\tilde{xad}_i) = xad_i + \epsilon_i^x - \epsilon_i^s + \epsilon_i^a, \tag{20}$$

where ϵ^a captures optimization error as well as any correlation between errors in advertising and revenue.

Recall now that both markup measures are constructed using $SALE$ in the numerator of their expressions, thus any measurement error in reported revenue appears in both markup measures. Then we can decompose both noise terms into the sum of ϵ^s and a measure specific noise term $\epsilon^{M'}$, that is for $M \in \{CA, PF\}$[52]:

$$\epsilon_i^M = \epsilon_i^s + \epsilon_i^{M'} \tag{21}$$

[50] Note that this definition of a signal-to-noise ratio—as well as the proceeding analysis—is robust to allowing also for a multiplicative and additive *bias* in log markups ie: $\log\left(\hat{\mu}_{M,i}\right) = \beta^M\left[\log\left(\mu_i^*\right) + \epsilon_i^M\right] + \alpha^M$ where β^M and α^M are constant within markup measures. For the sake of notational simplicity, this possible affine bias transformation will be suppressed. It is worth noting that the case of uninformative markups corresponds to $\beta^M = 0$, and $var\left(\epsilon^M\right) = \infty$, hence $SNR\left(\hat{\mu}_M\right) = 0$ while $\beta^M \epsilon^M$ retains a finite variance.

[51] Note that the subsequent analysis is unaffected by allowing for additive and multiplicative bias in $\log\left(\tilde{xad}\right)$. Such a bias could arise (for example) from optimization error.

[52] Then $\epsilon^{PF'}$ corresponds to the error on $ln(\hat{\theta}) - ln(COGS)$, while $\epsilon^{CA'}$ corresponds to the error on $-ln(COGS + PPEGT \times r)$

In addition to these errors—including the presence of the measurement error in revenue SALE on both sides of the equation–the other issue for the Dorfman-Steiner equation empirically is the unobserved advertising response elasticity, which could be correlated with the markup and the advertising share. For my purposes here, it is only necessary to assume that, while ε_{xad} may be correlated with the true underlying markup and true underlying advertising share, it is uncorrelated with the three errors listed above. Then, even when ε_{xad} is added to the equation, each of the measure-specific combined errors ($\epsilon^{M'}$) is uncorrelated with ε_{xad}. Then the following result follows:

Theorem 2 *Denote the squared correlation coefficient between* $-log\left(1-\tilde{xad}\right)$ *and* $log\left(\hat{\mu}^{M}\right)$ *by* $R^2_{X,M}$ *for* $M \in \{CA, PF\}$. *Then:*

$$\frac{R^2_{X,CA}}{R^2_{X,PF}} = \frac{1+\frac{1}{SNR(\hat{\mu}_{PF})}}{1+\frac{1}{SNR(\hat{\mu}_{CA})}}$$

The immediate take-away from theorem 2 is that if $log\left(\hat{\mu}_{CA}\right)$ is a better predictor of $-log\left(1-\tilde{xad}\right)$ than $log\left(\hat{\mu}_{PF}\right)$, CA markups must have a higher in-sample to signal-to-noise ratio, and vice versa. The intuition for this result is that all of the errors and biases affect both measures in the same way, contained in factors that cancel out when taking a ratio. Proof of theorem 2 is given in appendix E.2 and follows from this intuition.

The decision to compare to $-log\left(1-\tilde{xad}\right)$ may appear non-obvious at first, and indeed the result would follow with any monotonic transformation of $\tilde{xad}$. This particular form is motivated by a desire to preserve as much of the signal of $\log\left(\mu^*\right)$ as possible, motivated by the fact that equation (16) can be rewritten as:[53]

$$-log\left(1-\frac{\tilde{xad}}{\varepsilon_{xad}}\right) = \log\left(\mu^*\right)$$

[53] Additional motivation can be seen from taking a Taylor expansion of $-log\left(1-\tilde{xad}\right)$, which is given in appendix E.2.

5.3 Advertising Test: Results

Table 1: Advertising Tests

Dependent Variable:	$-log(1-xad)$							
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	0.0221***	0.0241***						
	(0.0008)	(0.0007)						
$log(\hat{\mu}_{CA})$	0.0276***		0.0247***		0.0271***		0.0251***	
	(0.0020)		(0.0022)		(0.0019)		(0.0022)	
$log(\hat{\mu}_{PF})$		0.0258***		0.0235***		0.0245***		0.0230***
		(0.0018)		(0.0020)		(0.0017)		(0.0020)
Fixed-effects								
Industry			Yes	Yes			Yes	Yes
Year					Yes	Yes	Yes	Yes
Fit statistics								
Observations	89,635	89,635	88,867	88,867	89,635	89,635	88,867	88,867
R^2	0.02012***	0.01972***	0.03346***	0.03327***	0.04817***	0.04708***	0.05965***	0.05875***
	(0.00093)	(0.00092)	(0.00119)	(0.00118)	(0.0014)	(0.00138)	(0.00154)	(0.00153)
Within R^2			0.01482	0.01464	0.01878	0.01767	0.01495	0.01402

Clustered (Firm level) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1. R^2 standard errors from Cohen et al. (2003).*

Notes: Regression of log markup measures on a function of the advertising share of revenue. Note that by theorem 2 the ratio of R^2 between measures gives the ratio of signal-to-noise ratios between two markup measures. The results then imply that signal-to-noise ratios are nearly identical, but CA is slightly higher in sample. Given the number of firms, this can be interpreted as largely driven by within sector × time as is shown by the inclusion of industry and year fixed effects

Table 1 presents the observation-level regression results linking markup measures to advertising shares. Both cost accounting (CA) and production function (PF) markups exhibit positive and significant coefficients, indicating that both measures provide a signal of underlying markups. The positive and statistically significant coefficients for both measures reject the null hypothesis that they are entirely noise. This confirms that both CA and PF measures capture meaningful signals of underlying markups, even in the presence of measurement error and attenuation bias. Additionally, CA regressions yield a slightly higher R^2, which by theorem (2) implies CA has a higher in-sample signal-to-noise ratio.

The first two columns denote the baseline regression without any controls. Columns 3–8 of Table 2

incorporate industry and year fixed effects, allowing for multiplicative bias in $\hat{\mu}_M$[54] to vary across these dimensions. This controls for unobserved sector- and time-specific factors that may influence the bias on a given markup, or advertising behavior. Including these fixed effects does not qualitatively change the story. The persistence of positive coefficients across these specifications reinforces the robustness of the relationship between markup measures and advertising shares. Additionally, this clarifies the result: both measures have similar signal-to-noise ratios within sector years.

Similar signal-to-noise ratios imply that disagreement between the two measures is the result of each measure deviating from underlying markups in approximately equal magnitude. It then stands to reason that averaging across measures may improve performance. Appendix E.3 includes a horse race regression including both measures in a single specification. The R^2 increases modestly from the CA-only specifications, and in all specifications, the CA receives a larger coefficient.

While similarity within sector-year is notable, this is likely in part driven by the fact that the production function component of the PF markup was estimated at the sector $\times$ time level. In other words, within-sector-year variation in PF is driven only by accounting data, driving the similarity with CA. But the two measures disagree more across sectors $\times$ time.

Table 2: Advertising Test: Sector $\times$ Time Averages

Dependent Variable:					$-log(1-xad)$			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	-0.0098***	0.0240***						
	(0.0037)	(0.0014)						
$log(\hat{\mu}_{CA})$	0.0279***		0.0282***		0.0283***		0.0285**	
	(0.0024)		(0.0085)		(0.0034)		(0.0100)	
$log(\hat{\mu}_{PF})$		0.0278***		0.0324***		0.0238***		0.0239**
		(0.0037)		(0.0107)		(0.0078)		(0.0105)
Fixed-effects								
Industry			Yes	Yes			Yes	Yes
time					Yes	Yes	Yes	Yes
Fit statistics								
Observations	1,030	1,030	983	983	1,030	1,030	983	983
R^2	0.12008	0.05255	0.32044	0.27843	0.25649	0.19283	0.43454	0.40117
Within R^2			0.09883	0.04313	0.11867	0.04321	0.08050	0.02623

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

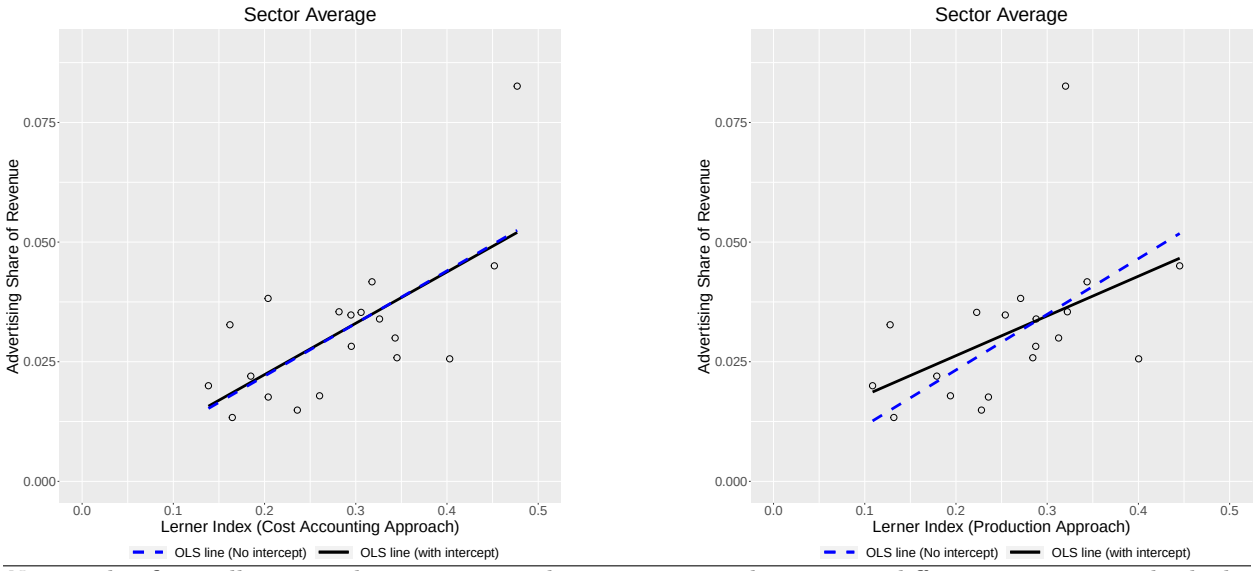
Notes: Regression of sector averages of log markup measures on a function of the advertising share of revenue. Note that by theorem 2 the ratio of R^2 between measures gives the ratio of signal-to-noise ratios between two markup measures. The results then imply that CA has a substantially higher signal-to-noise ratio in the sector $\times$ time cross-section.

To demonstrate this, Table 2 shifts to sector$\times$ year averages. All specifications yield much larger R^2's than in table 1, indicating that averaging significantly reduces the attenuation biases. In all specifications,

[54] Ie additive bias on $\log(\hat{\mu}_M)$

CA has a much higher R^2 than PF (often almost double), indicating a substantially larger signal-to-noise ratio in the cross-sector $\times$ time cross-section than in the within-sector $\times$ time cross-section. Given that sector $\times$ time is the level at which the production function component is estimated and allowed to vary, this result suggests that the estimated output elasticity provides relatively little signal about true underlying markups. This reinforces the result in section 4, which showed that SALE/COGS drove 94% of the variation in PF. Note that the non-capital share of variable costs (that is the other component of the CA markup besides SALE/COGS) is *allowed* to vary within sectors according to the firm-level data, though it varies much less within sectors than it does between sectors. Thus, because the unique component in PF does not vary within sectors, and the unique component to CA varies less within sectors, cross-sector comparisons best highlight the differences between the two measures.

Figure 7: Advertising Vs Markup Measures



Notes: This figure illustrates the cross-sectional variation in markups across different sectors using both the cost accounting and production function approaches. Note this is the dimension in which we see the highest disagreement between measures.

Additional insight can be drawn from Figure 7, which presents a scatter plot of sector-level averages of $\tilde{xad}$ and log markup measures from the two methods. Although the small number of observations limits the strength of conclusions, there is a clear alignment of the with- and without-intercept regression lines in the cost accounting (CA) plot, that does not appear in that of PF. The CA plot closely resembles the linear relationship suggested by (16) with constant ε_{xad}. Regression coefficients at the sector level, along with alternative specifications reflecting different assumptions about error terms, are provided in appendix E.3 for further robustness analysis.

6 Applications

Both the cost accounting (CA) and production function (PF) approaches provide useful markup estimates, but CA's simplicity and data-thriftiness offer some advantages. These observations motivate the exploration of specific applications where the CA approach's strengths are most evident. First, because CA does not require an assumption that firms share production functions, it is particularly suited to estimating markups for tech giants like Microsoft, Alphabet, and Meta. Second, its flexibility enables the study of how markups evolve across the firm life cycle, where production technologies and market conditions may change significantly. Finally, the CA approach is well-suited to capturing the cyclicality of markups in response to aggregate economic shocks, offering insights into macroeconomic dynamics and policy. Each of these applications underscores the practical advantages of the CA approach in addressing diverse and complex questions in the study of market power.

The cost accounting method is the only markup estimation method that can be implemented using only observation-specific information[55]. In principle it can be implemented with just 1 observation, attaining an estimate of a markup from just the accounting data of one firm in one year. In some extremely data-scarce environments, it would then be the only feasible option.

PF requires much more data. Estimation requires a large panel of firms that are assumed to have identical production technology. Even with COMPUSTAT, (the largest wildey available data set of US firms) parameters need to be estimated at the two-digit level in five-year rolling windows. This presents two weaknesses, the first is the required assumption that firms within sectors have identical technology parameters. Indeed there are contexts where it would be desirable to weaken this assumption. Additionally, CA relaxes the implicit restriction that technology does not vary over the firm life cycle, allowing for learning by doing. Lastly, CA does not require that any parameters be estimated within rolling windows, which could in principle dampen the measurement of business cycle effects.

Each of these observations motivates one of the following applications. The flexibility across firms motivates a case study of tech giants, without the need to assume that these companies share production technologies. The flexibility across the firm life cycle motivates a test of models that imply life cycle dynamics in markups. Lastly, the lack of rolling window estimation motivates an investigation of the response of markups to business cycle shocks. Collectively these applications highlight that wider use of the cost accounting approach opens additional avenues for investigations regarding market power.

[55] Conditional on the user cost of capital

6.1 Seven Magnificent Markups

In the wake of the Dotcom crash, several large technology companies rose to prominence. These "tech giants" built and retain large market shares in the information sector and hold market caps comprising a large portion of the *S&P*500. By 2013 the acronym *FANG* rose to prominence first in the financial punditry spheres with the likes of TV's Jim Crammer. This eventually became a household term to describe the collective of (at the time trading under the name) *Facebook*, *Amazon*, *Netflix*, and (at the time trading under the name) *Google*. Since then, the collection of firms considered *tech giants* has varied[56], but the size and importance of such firms has persisted.[57]

Large tech companies have a long history of drawing the ire of regulators and economic pundits alike.[58] In 1990 the Federal Trade Commission (FTC) launched an anti-trust investigation against Microsoft, and in 2000 Microsoft was found guilty of engaging in an illegal monopoly in the personal computer market. [59] Recently, a district court found that Google was acting as an illegal monopoly in the search engine market, with court documents revealing that it "enjoys an 89.2% share of the market for general search services, which increases to 94.9% on mobile devices".[60]

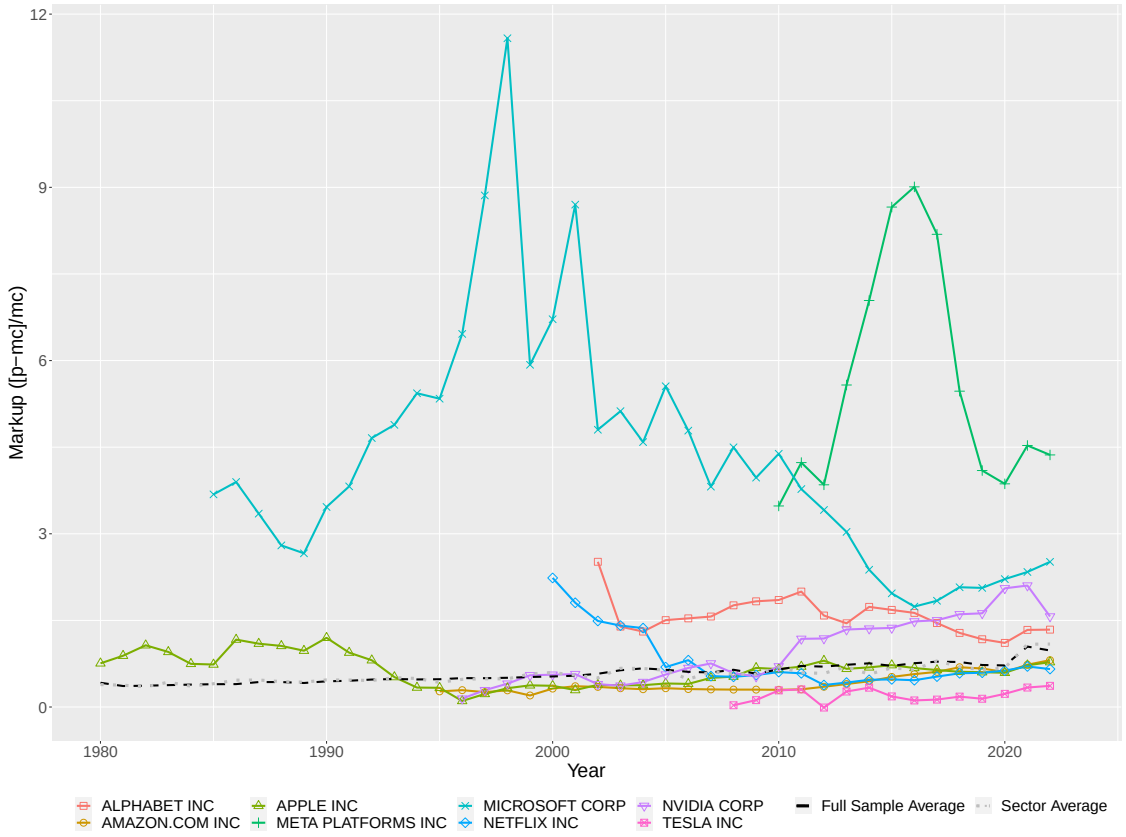
The cost accounting method offers a straightforward way to investigate the markups of these tech giants, using only accounting data from COMPUSTAT. Given that these companies are all publicly traded and founded after 1950, it is straightforward to calculate year-level markups for each firm from the year of their initial public offering (IPO) to the last year in our data set.

The production function method is not feasible at the firm level without grouping firms and assuming equal production function parameters between them. Given that all but one of the aforementioned companies are in the information sector[61] the typical approach of the production function method would mean assuming that all firms share production parameters. Implementation of an IO-style demand estimation approach would require demand-side data for each firm (and their direct competition) across each product and year considered. CA on the other hand can be implemented directly from the accounting data, without an assumption linking production technology.

Figure 8 presents the annual markups for the eight companies mentioned thus far in this subsection[62] from their IPO to 2022. For comparison, Full sample and sector averages are included in the plot. Microsoft, Meta, Alphabet, and Nvidia see markups consistently higher than these averages. There is significant heterogeneity

[56] And at the time of this paper *Magnificent Seven* is the term of choice to describe tech giants, consisting of Alphabet, Apple, Amazon, Tesla, Meta, Microsoft, Nvidia.
[57] Hobbs
[58] Payne
[59] v. Microsoft Corp [a]v. Microsoft Corp [b]
[60] Press
[61] with the exception of *Tesla* in the manufacturing sector, despite many annalists viewing it as more of a tech company.
[62] The seven members of the Magnificent Seven and Netflix.

Figure 8: Tech Giant Markups



*Notes:*This figure presents the annual markups for leading tech firms from their IPO to 2022, alongside the full sample and sector averages. Note the scale of the Y axis puts Microsoft in the 99.9 percentile at its peak.

in life cycle dynamics. Microsoft and Meta follow an inverted U path, entering with large markups, eventually rising to astronomical peaks before returning to levels similar to their peers. Netflix and Alphabet see slight downward trends following their IPOS. NVIDIA is unique in seeing a gradual but persistent markup growth, while markups for Apple and Amazon are relatively flat.

Microsoft sees a peak of its markups to $1,150\%$ of marginal cost in 1998, placing it well into the 99.9 percentile of markups in that year. In the years between the start of the FTC's investigation of the company in 1990, and its eventual conclusion Microsoft goes from a firm with a notably large markup relative to the rest of the sample to an outlier even compared to its closest contemporaries. Following (or slightly before) the eventual ruling the company sees a notable decline in markups, though they remain large relative to the other tech giants.

6.2 Markups Over the Firm Life Cycle

How do markups vary over the firm life cycle? There are several competing theories suggesting different patterns. The customer market literature argues that firms initially charge low markups to increase sales and raise markups once they have built market share and consumer awareness[63]. These models imply that markups rise over the firm life cycle.[64] On the other hand, the canonical Schumpeterian creative destructive framework predicts that competition increases the longer a firm is in the market, which would imply markups fall over the firm life cycle.

The production function approach is reliant on exploiting the changes in inputs for a given firm across years. It is then necessary to assume that production technology is constant throughout the firm life cycle. The cost accounting approach is not reliant on the assumption that firm technology is unchanged through the firm life cycle, and thus better fit for investigating this question.

Fitzgerald and Priolo [2018] utilizes a demand estimation approach to answer this question within the consumer food industry. They find no evidence of a trend in markups over the firm life cycle in this market. This is suggestive of a larger pattern, but different sectors may be more or less sensitive to any of the aforementioned mechanisms. The use of cost accounting markups allows for such an investigation across all sectors.

Firms enter COMPUSTAT data following their IPO, and it is from that point forward that we can observe a firm. This may dampen the effects argued for in the "customer markets" literature, as some firms may develop notable market shares before their IPO. Following an IPO firms receive an influx of capital, and thus expand afterward. If firms increase market shares by setting temporarily low markups, we should see a general rise in markups in the years following an IPO. Additionally, limiting observation to a firm's IPO does not rule out the ability to observe the effects of creative destruction, as the mechanism is more relevant to the latter end of the firm life cycle.

Figure 9 presents decadal averages of markups and market share[65] for IPOs and non-IPOs in the data set. IPOs see slightly higher markups than the non-IPOs with both following a similar upward trend over time. IPOs have market shares that are on average just a fraction of their incumbent counterparts. While an IPO does not represent a 'new firm' under the strictest definition, firms are relatively small (in terms of market share) at the time of their IPO, and grow later in their lives.

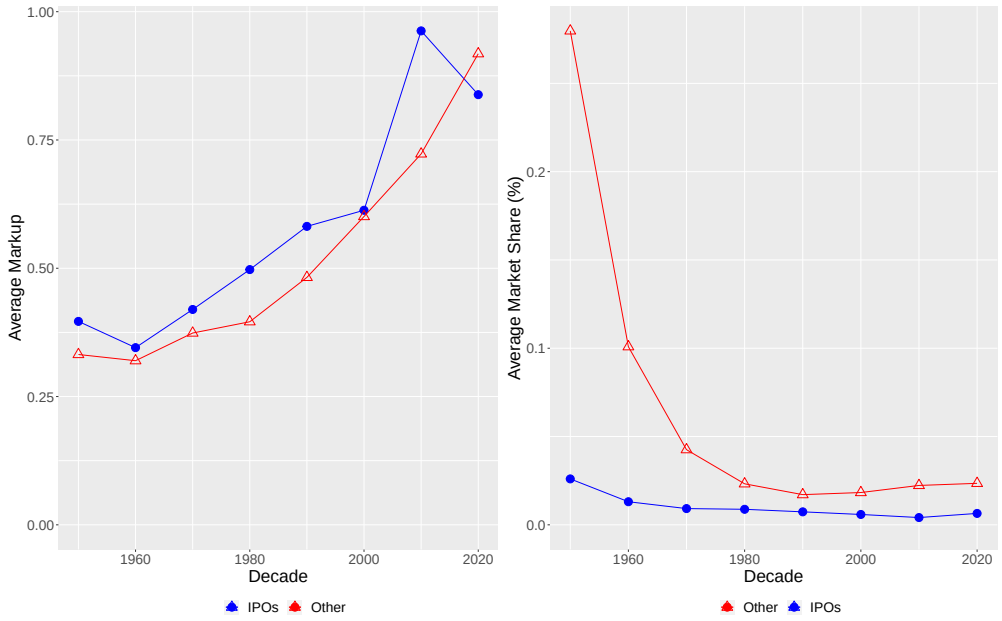
Figure 10 presents the unconditional average level of markups, and market share in the years following an IPO. Older firms on average have larger market share, with average market shares more than doubling

[63] See Fitzgerald and Priolo [2018] for a more in-depth literature review.

[64] A model with price-taking firms that increase their efficiency via learning by doing would also predict a rise in markups, as marginal costs may fall over time.

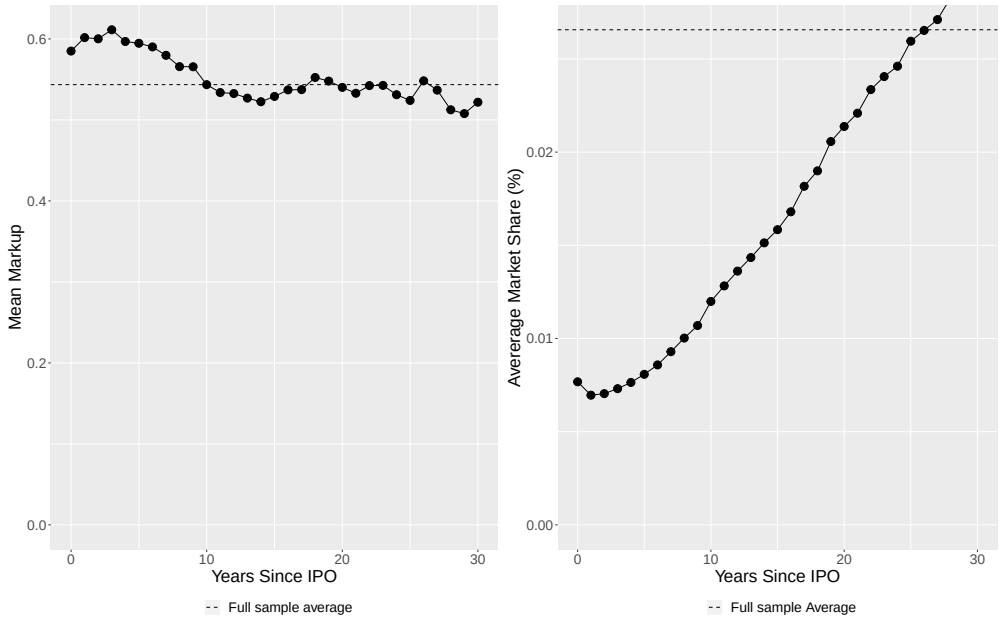
[65] Here market share is computer as the ratio of sales for a given firm in a given year, to the total sales of all firms in that year.

Figure 9: IPOs Over Time



Notes: This figure shows decadal average markups and market shares for IPO and non-IPO firms.

Figure 10: Life Cycle Dynamics: Unconditional Averages



Notes: The figure tracks the average markup and market share for firms at different stages after their IPO. Older firms on average have on average larger markup shares but lower markups.

within 15 years of IPO. Additionally, the average only crosses the full sample average after 25 years post-IPO. Despite this substantial rise in market shares, average market shares are nearly flat with a slight downward

trend.

Unconditional averages do not necessarily paint a clear picture of life cycle dynamics. They are influenced by the trend in average markups, changes in the age distribution across time, and survivorship bias. A simple linear regression is more than capable of controlling for these issues and estimating the life cycle component of markups directly. Including time-fixed effects directly addresses any concerns about underlying trends, as well as any issue with changes in the age distribution. Concerns about survivor-ship are addressed in two ways. First, the inclusion of firm fixed effects absorbs any issues related to the correlation between markups and time of exit. Second, an indicator for exit[66] absorbs any issues that can arise from patterns of markups near failure and the timing of exit. Then we are left with a regression of the form:

$$\mu_{it} - 1 = \vec{\beta} \cdot \mathbf{1}\left\{age\right\}_{it} + \alpha_i + \gamma_t + \alpha_{it}^{Exit} + \epsilon_{it} \tag{22}$$

Where α_i, γ_t, and α_{it}^{Exit}, are firm, age and exit fixed effects respectively, and $\mathbf{1}\left\{age\right\}_{it}$ is an age indicator for $age \in \{1, ..., 19, 20+\}$. Then $\hat{\vec{\beta}}_j$ gives an estimate of the conditional average life cycle component at age j.

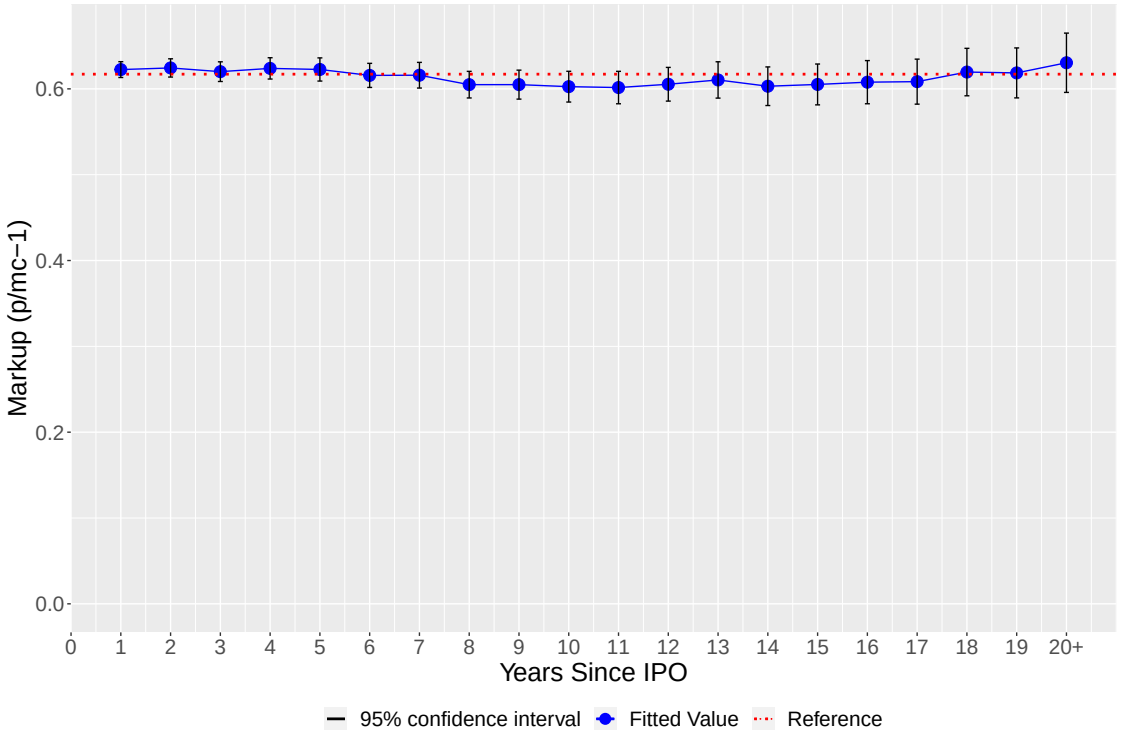
Figure 11 presents the fitted value from (22) for an average[67] firm of each age in the year 2018. The reference line represents the markup for an average firm in the year 2018. There is no clear trend in markups over the firm life cycle, with the gap between the lowest and highest fitted value being negligible compared to the overall distribution of markups. Standard error bars are also extremely tight. None of the estimates are statistically different from 0 or each other. We can confidently reject the existence of any trend in markups over the firm life cycle.

This result is surprising given the variety of mechanisms that would lead to various life cycle trends in markups. The effect of any such mechanism must therefore be small, or nonexistent. With respect to the consumer market framework, there is no evidence of firms setting low markups early into their life cycle to build market share. While this does not imply that no firms attempt a strategy of low markups to build market share, any such strategies do not bear out in the aggregate. We cannot fully rule out some variation of creative destruction leading to exit, but even removing the indicator for firm exit does not lead to any trend in markups either.

Alternative specifications and unconditional averages by sector and decade of entry are given in appendix F. There is no evidence of variation in markups over the firm life cycle in the alternative specifications once firm and time-fixed effects are included, and there is no clear life cycle trend for any of the two-digit NAICS industries. Appendix F also contains an estimate of the life cycle component of market shares estimated in the same manner as (22). There is a clear trend in market shares over the firm life cycle. Firms gain market

[66] Constructed as an indicator of a firm's last year in the sample
[67] Constructed by taking the average of the firm fixed effects

Figure 11: Markups over the firm life cycle: Conditional Averages



Notes: Fitted values pertain to an average firm in the year 2018, with 95% confidence intervals attained from standard errors clustered at the firm level. The reference line is constructed by taking the sum of the fixed effect for the year 2018, the average of all firm fixed effects, and the average of the age coefficients, and can thus be interpreted as the expected markup for an average firm in 2018.

share as they age, but there is no evidence that they are so by charging lower markups.

6.3 Markup Cyclicality

An important prediction of New Keynesian (NK) business cycle models is that of counter-cyclical markups. Models in the NK tradition predict a short-run decrease in markups in response to positive demand shocks. This effect drives the transmission mechanism of monetary and fiscal policy in these models. Markups above their long-run level lead to a negative output gap; ie real GDP below its long-run level. The 'stickiness' of business cycles in these models comes from a combination of 'nominal' (ie sticky prices or sticky wages) and 'real' rigidities, the latter of which emerges from strategic interactions between firms.[68] This leads to an insistent output gap, and motivates the role of monetary and fiscal policy. This role arises through the idea that with prices held fixed, a positive demand shock leads to inflationary pressure which raises marginal costs and lowers markups for all firms, addressing the nominal and real rigidities in one clean motion. This

[68] A firm adjusting its prices will not return to the long-run level when its competitors are charging markups above the long-run level, instead adjusting to a price that remains above the long-run level. See Ball and Romer [1990].

effect drives the transmission mechanism in canonical single-agent sticky price/sticky wage models, as well as heterogeneous Agent New Keynesian (HANK) models.[69]

Despite the importance, there is no consensus as to the cyclical behavior of markups in the data. There is evidence supporting countercyclical markups [70] as well evidence to the contrary[71]. As noted by Nekarada and Ramey [2020] this discrepancy is driven by differences in how markups are measured. To my knowledge, this is the first attempt to estimate the cyclicality of markups aggregated from firm-level estimates, rather than relying on an aggregate measure or proxy.

In practice, the production function approach is reliant on estimating parameters in rolling windows. This could obscure or otherwise dampen high-frequency movements. Additionally, a demand-based approach is simply infeasible at such an aggregate level. Limiting to aggregate measures opens up alternative measures of markup estimation. One approach can be thought of as an aggregate production function approach; considering the implied markup of a representative firm. In practice, the output elasticity of the aggregate firm can then be calibrated via the Hall et al. [1986] generalized sollow residual [72], or via the aggregate labor share of income.[73] Another approach is to infer cyclicality from a proxy, such as Hall [2012] and Bils and Kahn [2000][74]. Limiting to aggregate measures is appealing for several reasons: it dramatically reduces data needs, it is plausibly representative of the entire economy, and it generally allows for the use of quarterly data. There are however downsides to this approach: it is not clear what a representative firm is actually measuring[75], and relying on aggregates ties one's hands on how to aggregate up. Focusing on firm-level markups allows one to experiment with different methods of aggregation as well as look at individual sectors separately.

To facilitate comparison to the existing literature, I will estimate structural vector auto regressions matching the specifications in Nekarada and Ramey [2020], using aggregated cost accounting markups in place of their markup measure.[76] In particular, I follow their specification to identify supply shocks using an SVAR including the Fernald utilization adjusted TFP growth, log real GDP per capita, three-month treasury bill rates and log markups, with shocks identified via a Choleski decomposition.[77] For demand shocks I follow their specification for monetary policy shock identified from an SVAR including log real GDP per capita, log GDP deflator, log commodity prices, the federal funds rate, and log markups, identified via a Choleski

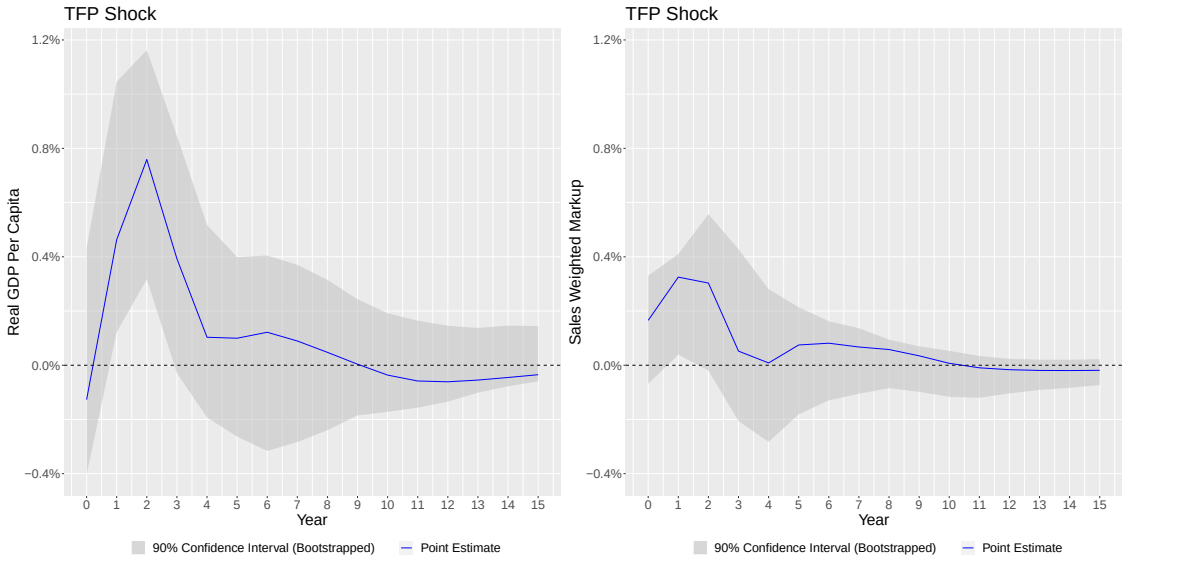
[69] See Smets and Wouters [2007] and Debortoli [2018]
[70] Ex: Galeotti and Schiantarelli [1998], Rotemberg and Woodford [1999], and Bils and Kahn [2000].
[71] Ex: Domowitz et al. [1986], Hall [2012] and Nekarada and Ramey [2020].
[72] See Haskel et al. [1995] and Marchetti [1999]
[73] See Rotemberg and Woodford [1999], Galeotti and Schiantarelli [1998], and Nekarada and Ramey [2020]
[74] There is also Domowitz et al. [1986]. which uses a sector-level cost accounting approach.
[75] NK models are reliant on heterogeneous firms.
[76] Note they utilize quarterly data, while this paper is limited to annual data
[77] Ordered as above. I include the years 1955-2018, removing the COVID period, given markups behave very erratically during this period

decomposition [78].

6.3.1 Markup Cyclicality: Results

Figure 12 measures markup cyclicality with respect to a supply shock. Plotted are the impulse responses on real GDP per capita and sales-weighted aggregate markups with respect to a one standard deviation shock in utilization-adjusted TFP, estimated as outlined above. NK models predict that markups should be *pro-cyclical* with respect to supply shocks. TFP shocks lower marginal costs, and price stickiness leads to an increase in markups In line with this prediction figure 12 admits pro-cyclical markups of moderate size.

Figure 12: Impulse Response: Supply Shock



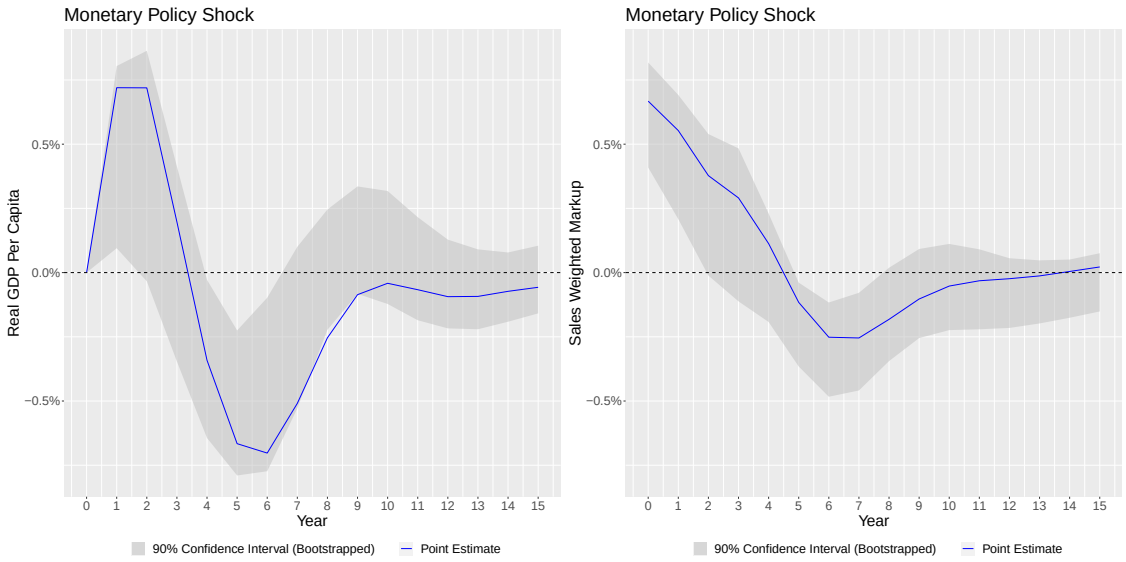
Notes: This figure shows the impulse response of real GDP per capita and sales-weighted aggregate markups to a one standard deviation TFP shock. 90% confidence intervals are constructed via bootstrapping with a block recursive structure. In line with New Keynesian models, markups are moderately pro-cyclical, rising in response to positive supply shocks.

Figure 13 measures markup cyclicality with respect to a demand shock. Plotted are the impulse responses on real GDP per capita and sales-weighted aggregate markups with respect to a one standard deviation shock in the federal funds rate, estimated as outlined above. The sign of the federal funds rate is swapped so that a positive shock coincides with a decrease in the federal funds rate. Figure 13 then gives the response to a positive monetary policy shock. NK models predict *counter-cyclical* markups, however figure 13 admits large and persistent pro-cyclical markups. This is not consistent with the transmission mechanism of NK models. [79]

The above results are based on sales-weighted aggregate markups across the entire economy. Appendix

[78] Ordered as above. I include the years 1955-2007, given the federal funds rate is uninformative in the following period.
[79] It is worth noting that the sample is limited to publicly traded firms. It may be the case that these firms admit qualitatively different cyclically than that of the economy overall.

Figure 13: Impulse Response: Demand Shock



Notes: This figure shows the impulse response of real GDP per capita and sales-weighted aggregate markups to a positive monetary policy shock. 90% confidence intervals are constructed via bootstrapping with a block recursive structure. Contrary to New Keynesian predictions, markups are highly pro-cyclical, suggesting a need to rethink the role of demand-side factors in driving markup fluctuations.

G includes the markup response to each shock broken down by sector. Wholesale firms see the highest degree of cyclically, while retail firms see the lowest. This suggests that nominal rigidities are highest for wholesale firms[80]. Additionally, we can compare how different different forms of aggregation affect estimates. To summarize measures of cyclically in a concise manner that allows for comparison across methods, I follow the norm in the literature and report the ratio of the areas below the impulse response functions in the first 5 years. The result can then be interpreted as the elasticity of markups with respect to real GDP, through each of the two channels.

Table 3 reports the elasticities through each channel for several forms of aggregation. For comparison, the main specifications of Nekarada and Ramey [2020] and predictions of Smets and Wouters [2007] are included. Alternative forms of weighting do not qualitatively change the behavior of cyclically. The direction of all results coincides with that of Nekarada and Ramey [2020], with markups pro-cyclical through both channels. This remains inconsistent with the main predictions of NK models. The results of Nekarada and Ramey [2020] suggest a higher degree of cyclically through supply shocks, while cost accounting suggests a higher degree of cyclically through demand shocks.

[80] In the case of price rigidity this could be driven by contracts with customers, and in the case of wage rigidity it could be the product of unionization.

Table 3: Impulse Response Elasticities

	Monetary policy	TFP
Cost Accounting Markups		
Sales Weighted μ	3.56	0.57
Mean μ	3.31	0.52
Median μ	4.44	0.38
Cost Weighted μ	4.07	0.49
Sales Weighted μ Manufacturing	3.99	0.71
Sales Weighted μ Retail	1.33	0.33
Sales Weighted μ Wholesale	4.69	1.45
Nekarada and Ramey (2020)		
CD, 1947–2017	.92	1.05
CD, 1964–2017	1.12	.51
CES, 1947–2017	1.11	2.74
CES, 1964–2017	1.28	2.27
NK Benchmark		
Smets and Wouters (2007)	-.51	.18

Notes: This table reports the elasticities of various markup measures in response to monetary policy and total factor productivity (TFP) shocks. They are constructed by taking the ratio of the area under the IRF of markups to that of TFP, for the first 5 years after a shock. Markups derived from the cost accounting approach exhibit a stronger pro-cyclical response to demand shocks, compared to supply shocks. Different forms of averaging make a minimal difference, but Wholesale sees significantly higher cyclicality. Neither the directions nor scale of monetary policy reactions are in line with the NK benchmark.

7 Conclusion

This paper makes the case for a revival of the cost accounting (CA) approach to markup estimation, a data-thrifty, highly transparent method. Despite its simplicity, the CA approach produces estimates that closely mirror those of the more complex production function (PF) method, which has become the standard in macroeconomic research. However, CA is not subject to the key criticisms PF is subject to and achieves a higher signal-to-noise ratio in predictions for advertising intensity across sectors and time.

A key substantive finding is that the rise in U.S. markups, as well as the rise of superstar firms, is robust to using CA, lending further credibility to this narrative, despite the critiques of PF. The component of the PF approach theoretically subject to biases does not drive its aggregate trend. To the contrary, 94% of the variation in the PF markup is due to its accounting component, which in turn is highly correlated with CA, but lacks a coherent micro-foundation.

In addition to fitting the Dorfman and Steiner [1954] advertising equation better, the simplicity and less stringent data requirements of the CA approach—in particular, its lack of need for a substantial panel for output elasticity estimation—also allow it to thrive in contexts where high-frequency markup movements are the focus of investigation.

Together, these results suggest that the CA method deserves broader application. At a more basic level,

the consistency of results across different methodologies, particularly regarding trends in market power, suggests that economists *can* meaningfully measure markups directly.

References

D. A. Ackerberg, K. Caves, and G. Frazer. Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451, 2015. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/43866416.

P. Aghion and P. Howitt. A model of growth through creative destruction. *Econometrica*, 60(2):323–351, 1992. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2951599.

P. Aghion, A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li. A theory of falling growth and rising rents. Working Paper 26448, National Bureau of Economic Research, November 2019. URL http://www.nber.org/papers/w26448.

K. Arrow. An extension of the basic theorems of classical welfare economics. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probabilit*, pages 507–532, 01 1951.

K. J. Arrow and G. Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3): 265–290, 1954. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1907353.

J. Asker, J. Farre-Mensa, and A. Ljungqvist. Corporate investment and stock market listing: A puzzle? *The Review of Financial Studies*, 28(2):342–390, 2015. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/24465702.

D. Autor, D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen. Concentrating on the fall of the labor share. *The American Economic Review*, 107(5):180–185, 2017a. ISSN 00028282. URL http://www.jstor.org/stable/44250385.

D. Autor, D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen. The fall of the labor share and the rise of superstar firms. Working Paper 23396, National Bureau of Economic Research, May 2017b. URL http://www.nber.org/papers/w23396.

K. Bagwell. Chapter 28 the economic analysis of advertising. volume 3 of *Handbook of Industrial Organization*, pages 1701–1844. Elsevier, 2007. doi: https://doi.org/10.1016/S1573-448X(06)03028-7. URL https://www.sciencedirect.com/science/article/pii/S1573448X06030287.

L. Ball and D. Romer. Real rigidities and the non-neutrality of money. *The Review of Economic Studies*, 57 (2):183–203, 1990. ISSN 00346527, 1467937X. URL http://www.jstor.org/stable/2297377.

S. Basu. Procyclical productivity: Increasing returns or cyclical utilization? *The Quarterly Journal of Economics*, 111(3):719–751, 1996. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/2946670.

S. Basu and M. S. Kimball. Cyclical productivity with unobserved input variation. Working Paper 5915, National Bureau of Economic Research, February 1997. URL http://www.nber.org/papers/w5915.

S. Basu, J. G. Fernald, and M. S. Kimball. Are technology improvements contractionary? *The American Economic Review*, 96(5):1418–1448, 2006. ISSN 00028282. URL http://www.jstor.org/stable/30034981.

S. Berry, J. Levinsohn, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2171802.

M. Bils and J. A. Kahn. What inventory behavior tells us about business cycles. *American Economic Review*, 90(3):458–481, June 2000. doi: 10.1257/aer.90.3.458. URL https://www.aeaweb.org/articles?id=10.1257/aer.90.3.458.

S. Bond, A. Hashemi, G. Kaplan, and P. Zoch. Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *Journal of Monetary Economics*, 121:1–14, 2021. ISSN 0304-3932. doi: https://doi.org/10.1016/j.jmoneco.2021.05.004. URL https://www.sciencedirect.com/science/article/pii/S0304393221000544.

R. Boucekkine, H. Latzer, and M. Parenti. Variable markups in the long-run: A generalization of preferences in growth models. *Journal of Mathematical Economics*, 68:80–86, 2017. ISSN 0304-4068. doi: https://doi.org/10.1016/j.jmateco.2016.11.005. URL https://www.sciencedirect.com/science/article/pii/S0304406816302701.

T. F. Bresnahan. Chapter 17 empirical studies of industries with market power. volume 2 of *Handbook of Industrial Organization*, pages 1011–1057. Elsevier, 1989. doi: https://doi.org/10.1016/S1573-448X(89)02005-4. URL https://www.sciencedirect.com/science/article/pii/S1573448X89020054.

E. H. Chamberlain. *The Theory of Monopolistic Competition*. Cambridge: Harvard Univ. Press, 1933.

S. J. Davis, J. Haltiwanger, R. Jarmin, J. Miranda, C. Foote, and Éva Nagypál. Volatility and dispersion in business growth rates: Publicly traded versus privately held firms [with comments and discussion]. *NBER Macroeconomics Annual*, 21:107–179, 2006. ISSN 08893365, 15372642. URL http://www.jstor.org/stable/25554954.

J. De Loecker and P. T. Scott. Estimating market power evidence from the us brewing industry. Working Paper 22957, National Bureau of Economic Research, December 2016. URL http://www.nber.org/papers/w22957.

J. De Loecker, J. Eeckhout, and G. Unger. The Rise of Market Power and the Macroeconomic Implications*. *The Quarterly Journal of Economics*, 135(2):561–644, 01 2020. ISSN 0033-5533. doi: 10.1093/qje/qjz041. URL https://doi.org/10.1093/qje/qjz041.

M. De Ridder. Market power and innovation in the intangible economy. *American Economic Review*, 114 (1):199–251, January 2024. doi: 10.1257/aer.20201079. URL https://www.aeaweb.org/articles?id=10.1257/aer.20201079.

D. Debortoli. Monetary policy with heterogeneous agents : Insights from tank models . 2018. URL https://api.semanticscholar.org/CorpusID:171084113.

A. Dixit and V. Norman. Advertising and welfare. *The Bell Journal of Economics*, 9(1):1–17, 1978. ISSN 0361915X. URL http://www.jstor.org/stable/3003609.

A. Dixit and V. D. Norman. Advertising and welfare: Reply. *The Bell Journal of Economics*, 10(2):728–729, 1979. ISSN 0361915X. URL http://www.jstor.org/stable/3003365.

I. Domowitz, R. G. Hubbard, and B. C. Petersen. Business cycles and the relationship between concentration and price-cost margins. *The RAND Journal of Economics*, 17(1):1–17, 1986. ISSN 07416261. URL http://www.jstor.org/stable/2555624.

R. Dorfman and P. O. Steiner. Optimal advertising and optimal quality. *The American Economic Review*, 44 (5):826–836, 1954. ISSN 00028282. URL http://www.jstor.org/stable/1807704.

H. Döpper, A. MacKay, N. Miller, and J. Stiebale. Rising markups and the role of consumer preferences. *SSRN Electronic Journal*, 01 2021. doi: 10.2139/ssrn.3939126.

M. W. L. Elsby, B. Hobijn, and A. Şahin. The decline of the u.s. labor share. *Brookings Papers on Economic Activity*, pages 1–52, 2013. ISSN 00072303, 15334465. URL http://www.jstor.org/stable/23723432.

F. Etro. Technological foundations for dynamic models with endogenous entry. *European Economic Review*, 128:103517, 2020. ISSN 0014-2921. doi: https://doi.org/10.1016/j.euroecorev.2020.103517. URL https://www.sciencedirect.com/science/article/pii/S0014292120301483.

D. Fitzgerald and A. Priolo. How do firms build market share? Working Paper 24794, National Bureau of Economic Research, July 2018. URL http://www.nber.org/papers/w24794.

L. S. Foster, J. C. Haltiwanger, and C. Tuttle. Rising markups or changing technology? Working Paper 30491, National Bureau of Economic Research, September 2022. URL http://www.nber.org/papers/w30491.

M. Galeotti and F. Schiantarelli. The cyclicality of markups in a model with adjustment costs: Econometric evidence for us industry. *Oxford Bulletin of Economics and Statistics*, 60(2):121–142, 1998. doi: https://doi.org/10.1111/1468-0084.00091. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0084.00091.

A. Gandhi, S. Navarro, and D. A. Rivers. On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016, 2020. doi: 10.1086/707736. URL https://doi.org/10.1086/707736.

R. J. Gordon. Output fluctuations and gradual price adjustment. *Journal of Economic Literature*, 19(2): 493–530, 1981. ISSN 00220515. URL http://www.jstor.org/stable/2724154.

G. M. Grossman and C. Shapiro. Informative advertising with differentiated products. *The Review of Economic Studies*, 51(1):63–81, 1984. ISSN 00346527, 1467937X. URL http://www.jstor.org/stable/2297705.

G. Gutiérrez and T. Philippon. Investment-less growth: An empirical investigation. Working Paper 22897, National Bureau of Economic Research, December 2016. URL http://www.nber.org/papers/w22897.

R. E. Hall. The relation between price and marginal cost in u.s. industry. *Journal of Political Economy*, 96 (5):921–947, 1988. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/1837241.

R. E. Hall. The cyclical response of advertising refutes counter-cyclical profit margins in favor of product-market frictions. Working Paper 18370, National Bureau of Economic Research, September 2012. URL http://www.nber.org/papers/w18370.

R. E. Hall, O. J. Blanchard, and R. G. Hubbard. Market structure and macroeconomic fluctuations. *Brookings Papers on Economic Activity*, 1986(2):285–338, 1986. ISSN 00072303, 15334465. URL http://www.jstor.org/stable/2534476.

J. Haskel, C. Martin, and I. Small. Price, marginal cost and the business cycle. *Oxford Bulletin of Economics and Statistics*, 57(1):25–39, 1995. doi: https://doi.org/10.1111/j.1468-0084.1995.tb00025.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0084.1995.tb00025.x.

B. Hobbs. Faang: An acronym that stands for five very successful tech companies that can move the stock market. *Buisness Insider*. URL https://www.businessinsider.com/personal-finance/

investing/what-is-faang#:~:text=FAANG%20actually%20began%20as%20FANG.%20The%20origin%20of,represent%20four%20tech%20stocks%20with%20outsized%20market%20appreciation.

L. Karabarbounis and B. Neiman. The global decline of the labor share. Working Paper 19136, National Bureau of Economic Research, June 2013. URL http://www.nber.org/papers/w19136.

M. S. Kimball. The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking*, 27(4):1241–1277, 1995. ISSN 00222879, 15384616. URL http://www.jstor.org/stable/2078048.

T. J. Klette and Z. Griliches. The inconsistency of common scale estimators when output prices are unobserved and endogenous. *Journal of Applied Econometrics*, 11(4):343–361, 1996. ISSN 08837252, 10991255. URL http://www.jstor.org/stable/2284929.

H. Latzer, K. Matsuyama, and M. Parenti. Reconsidering the Market Size Effect in Innovation and Growth. Working Papers ECARES 2019-31, ULB – Universite Libre de Bruxelles, Dec. 2019. URL https://ideas.repec.org/p/eca/wpaper/2013-298200.html.

R. Z. Lawrence. Recent declines in labor's share in us income: A preliminary neoclassical account. Working Paper 21296, National Bureau of Economic Research, June 2015. URL http://www.nber.org/papers/w21296.

J. D. Loecker and F. Warzynski. Markups and firm-level export status. *The American Economic Review*, 102 (6):2437–2471, 2012. ISSN 00028282. URL http://www.jstor.org/stable/41724661.

D. Marchetti. Markup and the business cycle: Evidence from italian manufacturing branches. 02 1999.

C. J. Nekarada and V. A. Ramey. The cyclical behavior of the price-cost markup. *Journal of Money, Credit and Banking*, 52(S2):319–353, 2020. doi: https://doi.org/10.1111/jmcb.12755. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jmcb.12755.

A. Nevo. Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 31(3):395–421, 2000. ISSN 07416261. URL http://www.jstor.org/stable/2600994.

S. Payne. Opinion today: What to do about big tech. *Finacial Times*. URL https://www.ft.com/content/a771c7ae-c99f-11e7-ab18-7a9fb7d6163e.

T. A. Press. Google loses massive antitrust case over its search dominance. *NPR*. URL https://www.npr.org/2024/08/05/nx-s1-5064624/google-justice-department-antitrust-search.

D. Raval. Testing the Production Approach to Markup Estimation. *The Review of Economic Studies*, 90(5): 2592–2611, 01 2023. ISSN 0034-6527. doi: 10.1093/restud/rdad002. URL https://doi.org/10.1093/restud/rdad002.

M. D. Ridder, B. Grassi, and G. Morzenti. The Hitchhiker's Guide to Markup Estimation. Discussion Papers 2210, Centre for Macroeconomics (CFM), July 2022. URL https://ideas.repec.org/p/cfm/wpaper/2210.html.

P. M. Romer. Endogenous technological change. *Journal of Political Economy*, 98(5):S71–S102, 1990. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/2937632.

J. N. Rosse. Estimating cost function parameters without using cost data: Illustrated methodology. *Econometrica*, 38(2):256–275, 1970. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1913008.

J. J. Rotemberg and M. Woodford. The cyclical behavior of prices and costs. Working Paper 6909, National Bureau of Economic Research, January 1999. URL http://www.nber.org/papers/w6909.

F. Smets and R. Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *The American Economic Review*, 97(3):586–606, 2007. ISSN 00028282. URL http://www.jstor.org/stable/30035013.

U. S. v. Microsoft Corp. U.s. v. microsoft: Court's findings of fact. *Justice.Gov Anti Trust Division*, a. URL https://www.justice.gov/atr/us-v-microsoft-courts-findings-fact.

U. S. v. Microsoft Corp. United states v. microsoft corp., 97 f. supp. 2d 59 (d.d.c. 2000). *Justia US Law*, b. URL https://law.justia.com/cases/federal/district-courts/FSupp2/97/59/2339529/.

E. Zhelobodko, S. Kokovin, M. Parenti, and J.-F. Thisse. Monopolistic competition: Beyond the constant elasticity of substitution. *Econometrica*, 80(6):2765–2784, 2012. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/23357240.

A Production Function Approach Derivation

Consider a production function of the form $Y = F(V, X^{-V})$ where V is perfectly flexible input. Then taking all other inputs as given (or at their optimal level), and given a price P^V for the variable input, the optimization sub-problem for the variable input V is summarized by the Lagrangian:

$$\mathcal{L}(V, \lambda) = P^V V - \lambda(F(V, X^{-V}) - Y^*)$$

λ is the key object of interest, it represents the shadow cost of an additional unit of output, ie the marginal cost, which can be seen directly from the envelope theorem. To back out the value of λ, first consider the first order condition for V:

$$\frac{\partial \mathcal{L}}{\partial V} = P^V - \lambda \frac{\partial F(V, X^{-V})}{\partial V} = 0 \tag{23}$$

Then (23) can be rearranged to solve for λ as:

$$\lambda = \frac{\partial V}{\partial F(V, X^{-V})} P^V = \frac{V P^V}{Y} \frac{1}{\theta_V}$$

Where $\theta_V \equiv \frac{\partial F(V, X^{-V})}{\partial V} \frac{V}{Y}$ is the output elasticity with respect to the variable input V. Then, given an output price of P, the markup ratio is given by:

$$\tilde{\mu}_{PF} = \frac{P}{\lambda} = \theta_V \frac{PY}{VP^V} \tag{24}$$

B Additional Figures

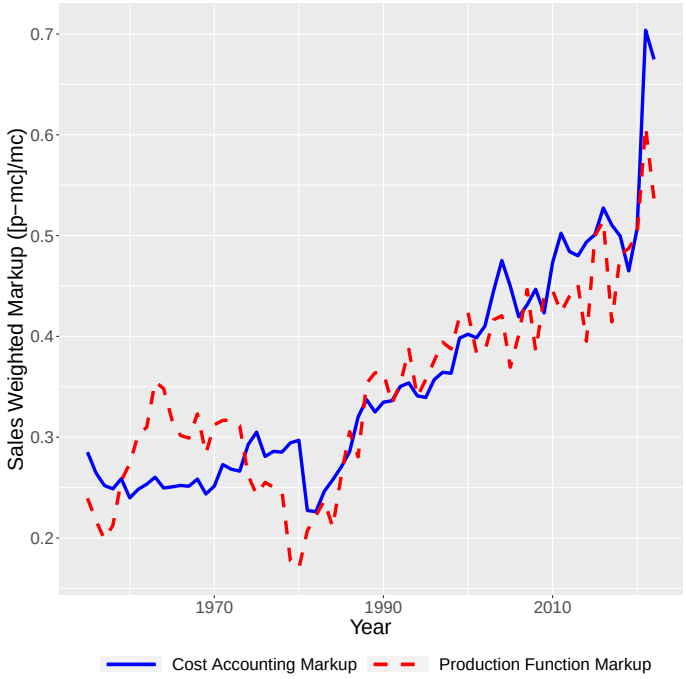


Figure 14: Markup Comparison: Sales-Weighted Markups

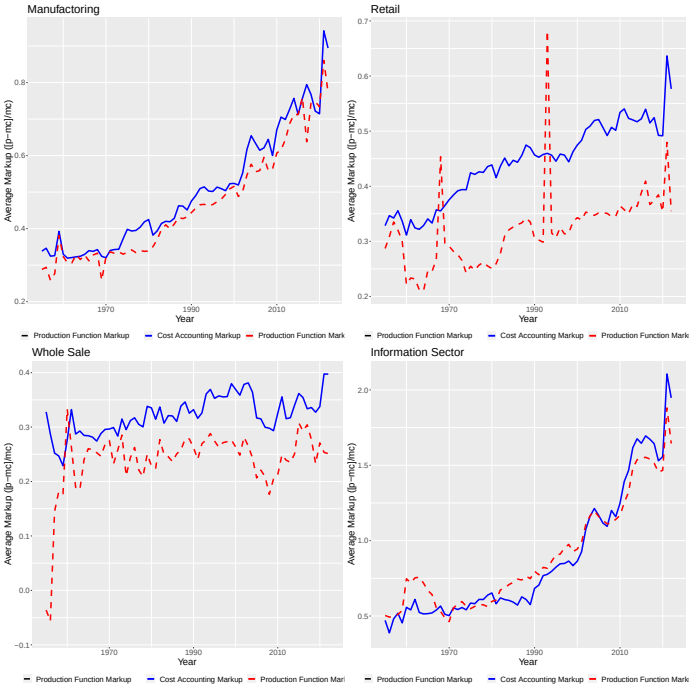


Figure 15: Markup Comparison: Sector level Annual Averages

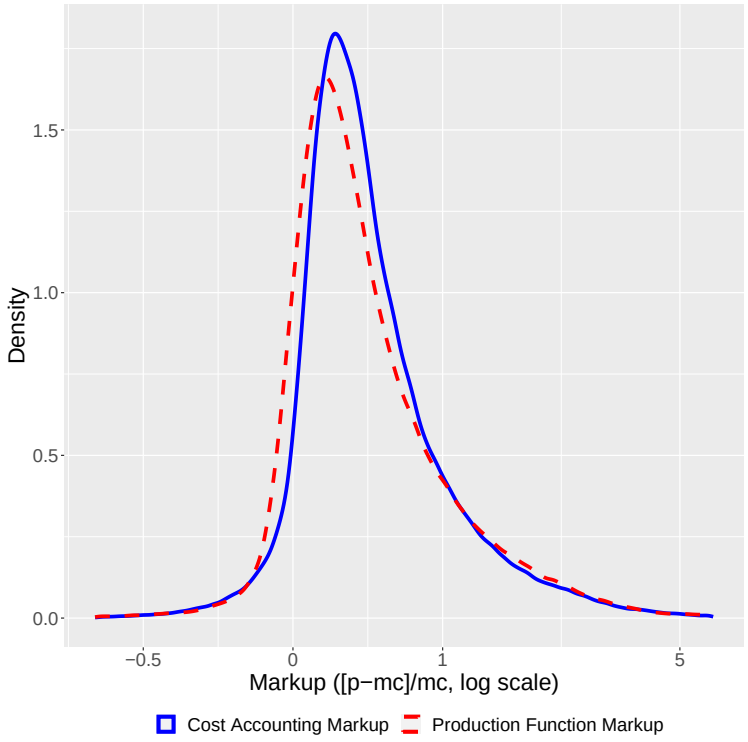


Figure 16: Markup Comparison: Full Density

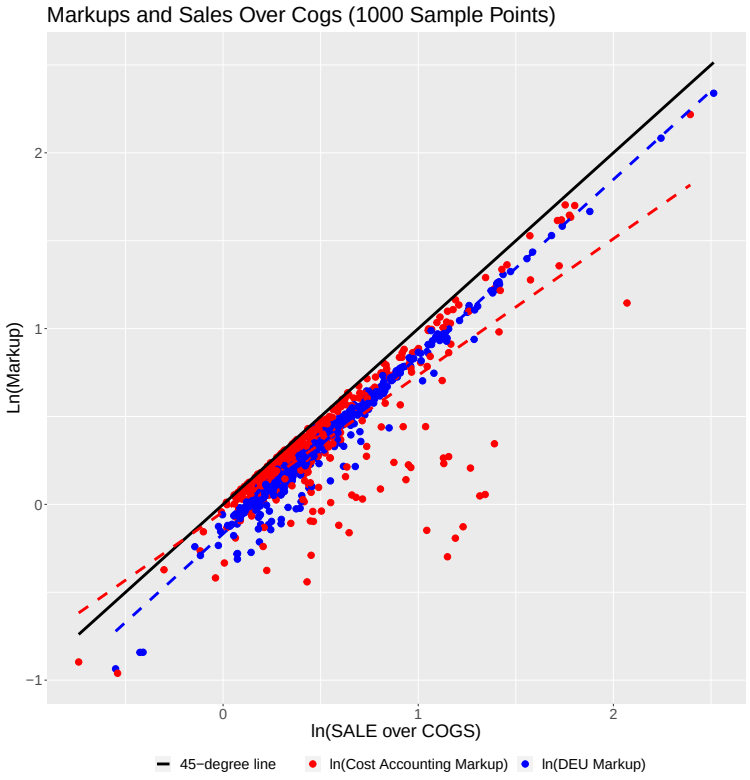


Figure 17: Markup Comparison: SALE over COGS

C Additional Markup Estimation Issues

C.1 Production Function Markups: Other Concerns

Many of the main concerns regarding the production function approach is noted in Bond et al. [2021]. In addition to the aforementioned observation of applicability of the Klette and Griliches [1996] critic, they also note that using a first stage to purge measurement error requires an AR(1) process on TFP, so while in principle one could follow the procedure allowing for a more complicated process on ω, assuming such a structure implies an additional structural error term. Furthermore Bond et al. [2021] demonstrates how any effects of quality of inputs onto demand leads to further bias. Additionally, Raval [2023] demonstrates that estimating markups from different inputs yields different markup estimates, treating capital, materials, and labor as alternative variable inputs, with each admitting different markup trends.

C.1.1 Production Function Markup Rule With Variable input utilization

Allowing for variable utilization and adjustment costs not only affects estimates of output elasticity but also affects marginal costs directly. To understand the implication on the production function markup rule, consider an alternative model with the following features: (i) firms face adjustment costs with respect to the variable input V given by the function $\zeta(V)$ which is dependent on some reference level V_r which is taken as given by the firms, (ii) firms have a choice over input utilization U^v, and (iii) variable input prices are a function of utilization, ie $P^V = P^v(U^v)$. In the simple case of labor as the variable input, we can imagine V denotes the number of employees and U^v is determined by either hourly adjustments or an increase in the intensity of work effort. Then we are left with production technology and cost function given as:

$$Y = F(U^v V; X) \tag{25}$$

$$C(UV; X) = P^v(U^v)V + \zeta(V) + \sum_i P_i X_i + F_C \tag{26}$$

Taking U^v and X as given at their optimal levels, the sub-problem for the variable input V can be represented with:

$$\mathcal{L}(V, \lambda) = P^v(U^v)V + \zeta(V) + \sum_i P_i X_i + \lambda(Y^* - F(U^v V; X))$$

From the first order condition with respect to V: $P(U) + \zeta'(V) = \lambda \frac{\partial Y}{\partial V} = \lambda \frac{Y}{V} \theta_V$, ie:

$$\frac{P}{\lambda} = \theta_V \left(\frac{PY}{V \left[P^v(U^v) + \zeta'(V) \right]} \right)$$

This then begs an important question: does $COGS$ include input adjustment costs? This is a difficult question to answer that remains out of the scope of this project, but the two most reasonable assumptions on the relation between $COGS$ and the true variable input are: (1) $COGS = VP^v(U^v)$ and (2) $COGS = V\left(P^v(U^v) + \zeta(V)\right)$. In the first case, we attain the true relation between output elasticity and the markup if and only if $\zeta' = 0$, and in the second case we attain the true relation if and only if $\zeta' = ae^V$ for some scalar a, and e denotes Euler's constant.

C.2 Cost Accounting: Other Concerns

In light of observing variable cost, there are three main concerns, the validity of accounting marginal, the measurement of user cost of capital, and the implications of variable utilization and adjustment costs.

The first issue comes down to a simple question of what is to be considered variable cost. On an accounting level, the division of accounting costs between COGS and XSGA is potentially motivated by industry norms, and not necessarily reflective of what costs are variable and which are fixed. Additionally, there is the question of fixed versus flexible capital. In the context of annual data, it would be unfair to assume capital stock fixed, though it is unclear that an assumption of full flexibility is much better[81].

The second issue arises from the fact that in the data, all costs are measured in units of dollars, however, accounting costs do not include the user cost of capital. That is, our measure of capital is more akin to a measure of the stock of capital than it is to its cost. As a result, we must infer capital costs by multiplying our measure of capital by a calibrated user cost of capital. Again siding on simplicity, I have chosen to adopt a simple approach to input the user cost of capital, though there is little evidence for the the value of deprecation and risk premium at an aggregate level. Advancement in measurement of this term is left as a future exercise and will be taken as a potential driver of measurement error.

To understand the final issue, let us again turn to the model with adjustment costs given by (26). Then from (1) we see that we properly observe variable cost only if $COGS = V\left(P^v(U^v) + \zeta(V)\right)$[82]. That is, if adjustment costs are reported as costs than the cost accounting approach properly measures variable costs, this could include things like the cost of hiring, or on-the-job training. However, if the costs are not reported directly or indirectly, then the cost accounting approach under states variable costs, such cases could include

[81] There is also the issue of rent, which is generally included in XSGA, though firm-owned property appears in PPEGT.
[82] which trivially includes the case $\zeta = 0$

cognitive costs of hiring, or office culture implications[83]

C.2.1 Cost Accounting Markups with Variable Input Utilization

The model with adjustment costs given by (26) also has implications on cost accounting markups through its effect on the degree of returns to scale on the variable cost (Γ). The effect of which depends on the specific function form of F, P^v, and ζ. For example if we take F as Cobb-Douglas, $\zeta(V) = \beta\left(V^z - V_r^z\right)$ and $P^v = P_r U^p$, then for a desired level of output Y^* marginal costs under cost minimization are given as[84]:

$$\lambda = Y^{*\left(\frac{1-\rho\theta_v+\sum_i \theta_i}{\rho\theta_v+\sum_i \theta_i}\right)}\left(\frac{\alpha^{\rho\theta_v}\prod_i P_i^{\theta_i}}{\Omega}\right)^{\left(\frac{1}{\rho\theta_v+\sum_i \theta_i}\right)}$$

Where for simplicity of notation $\rho \equiv \frac{p(z-1)}{p+z-1} + 1$ and $\alpha \equiv pP_r^{\left(\frac{z-1}{p+z-1}\right)}\left(\frac{z\beta}{p-1}\right)^{\frac{p}{p+z-1}}$. Then the assumption of constant marginal cost requires $1 = \rho\theta_V + \sum_i \theta_i$ which is quite a knife edge, and in the general district from that of constant returns to scale, ie $1 = \theta_V + \sum_i \theta_i$.

C.2.2 Proof

Consider a firm with production technology and cost function given as:

$$\begin{aligned} Y =& F(UV; X) \\ C(UV; X) =& P(U)V + \zeta(V) + \sum_i P_i X_i \end{aligned}$$

Let $\bar{V} = UV$; $\bar{P} = P(U)V + \zeta(V)$. Then the cost minimization problem can be decomposed into two subproblems. The first sub-problem takes U, V given function of $\bar{V}$, and minimizing costs given $Y = F(\bar{V}; X)$. The second sub-problem takes $\bar{V}$ as given and minimizes $\bar{P}$. The first sub-problem can be represented as:

$$\mathcal{L}(\bar{V}, X, \lambda) = \bar{P}(\bar{V}) + \sum_i P_i X_i + \lambda(Y^* - F(\bar{V}; X))$$

FOC:

$$\begin{aligned} \bar{V}: \quad & \bar{P}'(\bar{V}) = \lambda F_{\bar{V}}(\bar{V}, X) \\ X_i: \quad & P_i = \lambda F_{\bar{X}_i}(\bar{V}, X) \end{aligned}$$

[83] Unless employees recieve additional compensation for these burdens.
[84] Derivation given below.

With a Cobb-Douglas structure: $Y = \Omega \bar{V}^{\theta_V} \prod_i X_i^{\theta_i}$:

$$\begin{aligned} \bar{V}: \quad & \bar{P}'(\bar{V}) = \lambda \theta_V \frac{Y^*}{\bar{V}} \\ X_i: \quad & P_i = \lambda \theta_i \frac{Y^*}{X_i} \end{aligned}$$

The second sub-problem can be represented with:

$$\mathcal{L}(U, V, \lambda_2) = P(U)V + \zeta(V) + \lambda_2 \left[\bar{V} - UV\right]$$

FOC:

$$\begin{aligned} U: \quad & P'(U)V = \lambda_2 V \\ V: \quad & P(U) + \zeta'(V) = \lambda_2 U \end{aligned}$$

Then we have:

$$\bar{P}'(\bar{V}) = P(U) + \zeta'(V) = UP'(U)$$

Next take $\zeta(V) = \beta\left(V^z - V_r^z\right)$ and $P(U) = P_r U^p$. Then we can use the above expression alongside $UV = \bar{V}$ to show:

$$\begin{aligned} \zeta'(V) =& P'(U)U - P(U) \\ z\beta V^{z-1} =& pP_rU^p - P_rU^p \\ =& (p-1)P_rU^p \Longrightarrow \\ V^{z-1} =& U^p \left(\frac{p-1}{z}\right)\left(\frac{P_r}{\beta}\right) \Longrightarrow \\ U =& V^{\left(\frac{z-1}{p}\right)} \left[\left(\frac{z}{p-1}\right)\left(\frac{\beta}{P_r}\right)\right]^{\frac{1}{p}} \\ \bar{V} =& VU = V^{\left(\frac{P+z-1}{p}\right)} \left[\left(\frac{z}{p-1}\right)\left(\frac{\beta}{P_r}\right)\right]^{\frac{1}{p}} \Longrightarrow \\ V =& \bar{V}^{\left(\frac{p}{p+z-1}\right)} \left[\left(\frac{z}{p-1}\right)\left(\frac{\beta}{P_r}\right)\right]^{-\frac{1}{p+z-1}} \\ U =& \frac{\bar{V}}{V} = \bar{V}^{\left(\frac{z-1}{p+z-1}\right)} \left[\left(\frac{z}{p-1}\right)\left(\frac{\beta}{P_r}\right)\right]^{\frac{1}{p+z-1}} \end{aligned}$$

Then:

$$\begin{aligned}
\bar{P}'(\bar{V}) =& UP'(U) \\
=& pP_r\left(U\right)^p \\
=& pP_r\left(\bar{V}^{\left(\frac{z-1}{p+z-1}\right)}\left[\left(\frac{z}{p-1}\right)\left(\frac{\beta}{P_r}\right)\right]^{\frac{1}{p+z-1}}\right)^p \\
=& \bar{V}^{\left(\frac{p(z-1)}{p+z-1}\right)}pP_r^{\left(\frac{z-1}{p+z-1}\right)}\left(\frac{z\beta}{p-1}\right)^{\frac{p}{p+z-1}} \\
\equiv& \alpha\bar{V}^{\frac{1}{\rho}-1}
\end{aligned}$$

Where for simplicity of notation $\rho \equiv \frac{p+z-1}{p(z-1)}+1$ and $\alpha \equiv pP_r^{\left(\frac{z-1}{p+z-1}\right)}\left(\frac{z\beta}{p-1}\right)^{\frac{p}{p+z-1}}$. Then combine the above statement with the first-order condition from the first sub-problem:

$$\begin{aligned}
\alpha\bar{V}^{\frac{1}{\rho}-1} =& \lambda\theta_V\frac{Y^*}{\bar{V}} \implies \\
\bar{V} =& \left(\frac{\theta_V}{\alpha}\lambda Y^*\right)^{\rho}
\end{aligned}$$

Combing this and the other FOCs from the original sub-problem with the constraints:

$$\begin{aligned}
Y^* =& \Omega\bar{V}^{\theta_V}\prod_i X_i^{\theta_i} \\
=& \Omega\left(\frac{\theta_V}{\alpha}\lambda Y^*\right)^{\rho\theta_V}\prod_i\left(\frac{\lambda\theta_i Y^*}{P_i}\right)^{\theta_i} \\
=& \frac{Y^{*\left(\rho\theta_V+\sum_i\theta_i\right)}\lambda^{\left(\rho\theta_V+\sum_i\theta_i\right)}\Omega}{\alpha^{\rho\theta_V}\prod_i P_i^{\theta_i}} \implies \\
\lambda =& Y^{*\left(\frac{1-\rho\theta_V+\sum_i\theta_i}{\rho\theta_V+\sum_i\theta_i}\right)}\left(\frac{\alpha^{\rho\theta_V}\prod_i P_i^{\theta_i}}{\Omega}\right)^{\left(\frac{1}{\rho\theta_V+\sum_i\theta_i}\right)}
\end{aligned}$$

D The Role of SALE/COGS: Results

Table 4: Markup Decomposition

	$ln(\mu_{PF})$	...	...	$ln(\mu_{CA})$	...	...
(Intercept)	−0.152***	0.438***	0.262***	−0.043***	0.399***	0.408***
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
$ln\left(\frac{SALE}{COGS}\right)$	0.973***			0.798***		
	(0.000)			(0.001)		
$ln(\theta)$		0.561***			0.205***	
		(0.008)			(0.008)	
$ln\left(\frac{COGS}{COGS+r\times PPEGT}\right)$			−0.570***			0.294***
			(0.004)			(0.003)
N	254 662	254 662	254 662	254 662	254 662	254 662
R2	0.940	0.019	0.092	0.722	0.003	0.028

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Notes: This table presents regressions of key accounting components on production function (PF) and cost accounting (CA) markups. The R^2 in columns 1 and 2 present the explanatory power of the accounting and production function components on the PF. Note that the accounting component alone explains 94% of the variation in PF. The R^2 of columns 4 and 6 give the explanatory power of the PF accounting component and residual component on CA. Columns 3 and 5 present the explanatory power of components not used in the construction of each measure for comparison.

E Advertising

E.1 Additional Summary Statistics

Table (3) provides summary statistics of COMPUSTAT data across the sample from 1955-2023, measured in thousands of USD and deflated to 2010 dollars using GDP deflator. The sample used in the previous section, labeled Sample A includes all observations with data on COGS, PPEGT, and SALE (as well as the aforementioned trimming).

Sample B further limits to firms that report non-negative values for XAD, so that (4) and (5) can be estimated, and trimmed by 1% at the top and bottom of the ratio of XAD/SALES. The advertising balanced sample contains observations for 18 of the 19 two-digit NAICS industries, losing only "Public Administration" which consists of just 2 observations in the full sample.

Firms in the balanced sample have slightly higher revenue, approximately equal levels of variable costs, a higher level of fixed costs, and a slightly lower capital stock. Furthermore, observations in both samples have similar measures of "Years since IPO" and "Years to exit" which denote how long the firm has been in the sample, and how much longer it remains in the sample respectively. Differences in Sales, fixed costs and level of capital can potentially driven in part by firms who engage more heavily in advertising being more inclined to report advertising expenditure, as firms with larger markups should optimally engage in a higher

		Full Sample		
	Accounting Variable	Mean	Median	No. of Obs
Years Since IPO		11.95	8	257,808
Years to Exit		11.98	8	262,546
Cost of Goods Sold	COGS	1,453,652	93,523	280,513
SG&A	XSGA	387,750	32,618	280,513
Capital	PPEGT	1,641,505	63,064	280,513
Revenue	SALE	2,174,914	159,549	280,513
Advertising	XAD	71,939	2,593	97,521
Advertising Share	xad	0.04	0.02	97,521
Cost Accounting Markup	$\tilde{\mu} - 1$	0.66	0.34	278,704
		Advertising Balanced Sample		
	Accounting Variable	Mean	Median	No. of Obs
Years Since IPO		12.1	9	89,146
Years to Exit		10.94	8	89,654
Cost of Goods Sold	COGS	1,557,160	93,418	96,098
SG&A	XSGA	580,151	48,929	96,098
Capital	PPEGT	1,571,614	50,496	96,098
Revenue	SALE	2,526,687	173,340	96,098
Advertising	XAD	71,665	2,589	96,098
Advertising Share	xad	0.03	0.02	96,098
Cost Accounting Markup	$\tilde{\mu} - 1$	0.89	0.48	96,098

Table 5: Summary Statistics

level of marketing[85]. For this reason, the measure of advertising response elasticity can be interpreted as the elasticity among firms who engage in a level of advertising that they deem notable.

Figure 18 presents the distribution of the advertising share among the firms that report advertising. It is worth noting that prior to trimming, just 1.3% of firms that report advertising report a level of precisely 0, which lends further credence to the interpretation that firms who report advertising are firms who engage in a higher level of advertising.

[85]Note that for the relation with XSGA, advertising is mechanically a component of SGA, even if we assume firms who do not report XAD engage in 0 advertising, this discrepancy is not large enough to fully explain the difference.

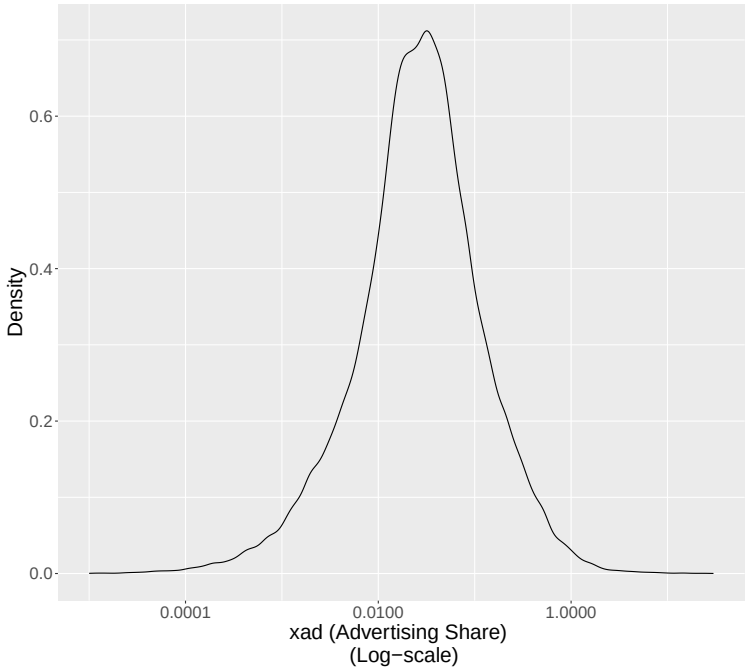


Figure 18: Advertising Share Density

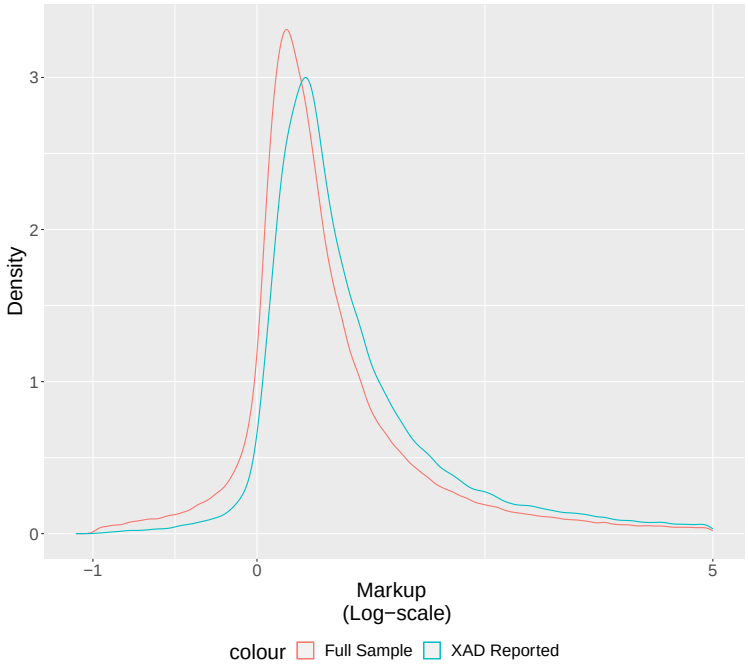


Figure 19: Markup Density

Figure 19 presents the distribution of markups as measured by the cost accounting approach, separate between firms who report advertising and the full sample. Conditioning on firms that report markups shifts the distribution to the right but does not seem to have a large effect on the overall shape of the

distribution. The rightward shift is consistent with the idea that firms reporting XAD to engage in higher levels of advertising, since by Theorem 1 we expect these to be firms with higher markups.

E.2 Proof of Theorem 2

Before showing the result let's first motivate the use of $-\log\left(1-\tilde{xad}\right)$. Combining (20) and (16):

$$\frac{1}{1-\tilde{xad}_i} = \frac{1}{1-\left(1-\frac{1}{\mu_i}\right)\eta_{xad,i}e^{\epsilon_i^a}e^{\epsilon_i^x}e^{-\epsilon_i^s}}$$

Then taking logs and taylor expanding around around $\eta_{xad,i}e^{\epsilon_i^a}e^{\epsilon_i^x}e^{-\epsilon_i^s}=1$:

$$\begin{aligned}
-log\left(1-\tilde{xad}_i\right) =& \log\left(\frac{1}{1-\left(1-\frac{1}{\mu_i}\right)\eta_{xad,i}e^{\epsilon_i^a}e^{\epsilon_i^x}e^{-\epsilon_i^s}}\right) \\
=& \log(\mu_i) + \left(\frac{\eta_{xad,i}e^{\epsilon_i^a}e^{\epsilon_i^x}e^{-\epsilon_i^s}-1}{\eta_{xad,i}e^{\epsilon_i^a}e^{\epsilon_i^x}e^{-\epsilon_i^s}\left(\mu_i-\eta_{xad,i}e^{\epsilon_i^a}e^{\epsilon_i^x}e^{-\epsilon_i^s}\right)}\right) + o\left(\left|\left|\eta_{xad,i}e^{\epsilon_i^a}e^{\epsilon_i^x}e^{-\epsilon_i^s}-1\right|\right|^2\right) \\
\equiv& log(\mu_i)+\epsilon_i^{xad} \qquad (27)
\end{aligned}$$

Now note that ϵ_i^{xad} is unlikely to be mean 0, and is correlated with μ, η_{xad}, ϵ^a, ϵ^x and ϵ^s by construction, but uncorrelated with ϵ^{CA} and ϵ^{PF}. Then by (18), (21) and (27) note that the square correlation coefficient between $log\left(\hat{\mu}_i^m\right)$ and $-log\left(1-\tilde{xad}\right)$ is given by:

$$\begin{aligned}
R^2_{X,M} \equiv& \frac{cov\left(-log\left(1-\tilde{xad}\right), log\left(\hat{\mu}^m\right)\right)^2}{\left[var\left(-log\left(1-\tilde{xad}\right)\right)\right]\left[var\left(log\left(\hat{\mu}_i^m\right)\right)\right]} \\
=& \frac{\left[cov\left(\epsilon^{xad}, log(\mu)\right)+cov\left(\epsilon^{xad},\epsilon^s\right)\right]^2}{\left[var\left(log(\mu)+\epsilon^{xad}\right)\right]\left[var\left(log(\mu)\right)+var(\epsilon^s+\epsilon^m)\right]} \\
=& \left(\frac{1}{1+\frac{var(\epsilon^s+\epsilon^m)}{var(log(\mu))}}\right)\frac{\left[cov\left(log(\mu)+\epsilon^{xad}, log(\mu)\right)+cov\left(\epsilon^{xad},\epsilon^s\right)\right]^2}{\left[var\left(log(\mu)+\epsilon^{xad}\right)\right]\left[var\left(log(\mu)\right)\right]} \\
=& \left(\frac{1}{1+\frac{1}{SNR(\hat{\mu}^m)}}\right)\frac{\left[cov\left(log(\mu)+\epsilon^{xad}, log(\mu)\right)+cov\left(log(\mu)+\epsilon^{xad},\epsilon^s\right)\right]^2}{\left[var\left(log(\mu)+\epsilon^{xad}\right)\right]\left[var\left(log(\mu)\right)\right]} \qquad (28)
\end{aligned}$$

Thus:

$$\frac{R^2_{X,CA}}{R^2_{X,PF}} = \frac{1+\frac{1}{SNR(\hat{\mu}^{PF})}}{1+\frac{1}{SNR(\hat{\mu}^{CA})}}$$

Note that squared correlation coefficients are unaffected by linear transformation. Thus the result holds when

adding constant within measure bias terms β^m and α^m:

$$log\left(\hat{\mu}_i^m\right) = \beta^m\left(log\left(\mu_i\right) + \epsilon_i^s + \epsilon_i^m\right) + \alpha^m$$

With the above $SR(\hat{\mu}^m) = \frac{var(log(\mu))}{var(\epsilon^s + \epsilon^m)}$ still denotes the signal to noise ratio, attain a value of 0 in the limit as $\beta^m \to 0$ while holding $\beta_m\left(var(\epsilon^s + \epsilon^m)\right)$ constant. Lastly from (28) note that as $var(\epsilon^s + \epsilon^m) \to \infty$, $R^2_{X,M} \to 0$, thus $R^2 \neq 0$ serves as a direct test of relevancy of markup measure m.

E.3 Advertising Test: Robustness

Dependent Variables:	$-log(1-xad)$				xad			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	-0.0519**	0.0102			0.0026	0.0242***		
	(0.0206)	(0.0108)			(0.0090)	(0.0076)		
$log\left(\hat{\mu}_{CA}\right)$	0.0563***		0.0231***					
	(0.0133)		(0.0024)					
$log\left(\hat{\mu}_{PF}\right)$		0.0767**		0.1049***				
		(0.0323)		(0.0118)				
$1-\frac{1}{\hat{\mu}_{CA}}$					0.1031***		0.1113***	
					(0.0296)		(0.0089)	
$1-\frac{1}{\hat{\mu}_{PF}}$						0.0350		0.1180***
						(0.0294)		(0.0164)
Fit statistics								
Observations	16	16	16	16	19	19	19	19
R^2	0.56134	0.28644	0.31915	0.19091	0.41586	0.07683	0.38037	-0.55070
Adjusted R^2	0.53001	0.23547	0.31915	0.19091	0.38150	0.02252	0.38037	-0.55070

IID standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6: Advertising Test: Horse Race

Dependent Variables:	xad				$log(xad)$			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	0.0208***	0.0265***			-3.750***	-3.747***		
	(0.0008)	(0.0006)			(0.0299)	(0.0264)		
$1 - \frac{1}{\mu_{CA}}$	0.0351***		0.0320***					
	(0.0027)		(0.0028)					
$1 - \frac{1}{\mu_{PF}}$		0.0199***		0.0245***				
		(0.0017)		(0.0021)				
$log\left(1 - \frac{1}{\mu_{CA}}\right)$					0.3658***		0.4421***	
					(0.0208)		(0.0201)	
$log\left(1 - \frac{1}{\mu_{PF}}\right)$						0.3105***		0.3387***
						(0.0151)		(0.0146)
Fixed-effects								
Industry			Yes	Yes			Yes	Yes
time			Yes	Yes			Yes	Yes
Fit statistics								
Observations	89,837	89,837	88,968	88,968	85,977	82,753	85,187	82,136
R^2	0.02023	0.01289	0.06867	0.06670	0.03263	0.03743	0.14789	0.14067
Within R^2			0.01525	0.01316			0.04470	0.04317

Clustered (Firm level) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 7: Advertising Test: Alternative Specifications

F Markups Over the Firm Life Cycle

Dependent Variable:	Markup						
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Constant	0.6002***	0.1369***					
	(0.0057)	(0.0108)					
Firm Age	-0.0034***	-0.0077***	-0.0078***	-0.0061***	6.89×10^{-5}		
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)		
Time		0.0121***					
		(0.0003)					
$Age \times I[Age < 6]$						0.0003	8.48×10^{-5}
						(0.0010)	(0.0010)
$Age \times I[5 < Age < 11]$						-0.0015**	-0.0018***
						(0.0007)	(0.0007)
$Age \times I[10 < Age < 20]$						-0.0008*	-0.0010**
						(0.0005)	(0.0005)
$Age \times i[Age > 20]$						-3.6×10^{-5}	-0.0002
						(0.0004)	(0.0004)
Fixed-effects							
Decade			Yes	Yes	Yes	Yes	Yes
Industry				Yes			
Firm					Yes	Yes	Yes
Exit indicator							Yes
Fit statistics							
Observations	234,968	234,968	234,968	232,161	234,968	234,968	234,968
R^2	0.00268	0.05231	0.05511	0.15457	0.78440	0.78445	0.78447
Within R^2			0.01357	0.00897	2.66×10^{-7}	0.00026	0.00036

Clustered (Firm) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

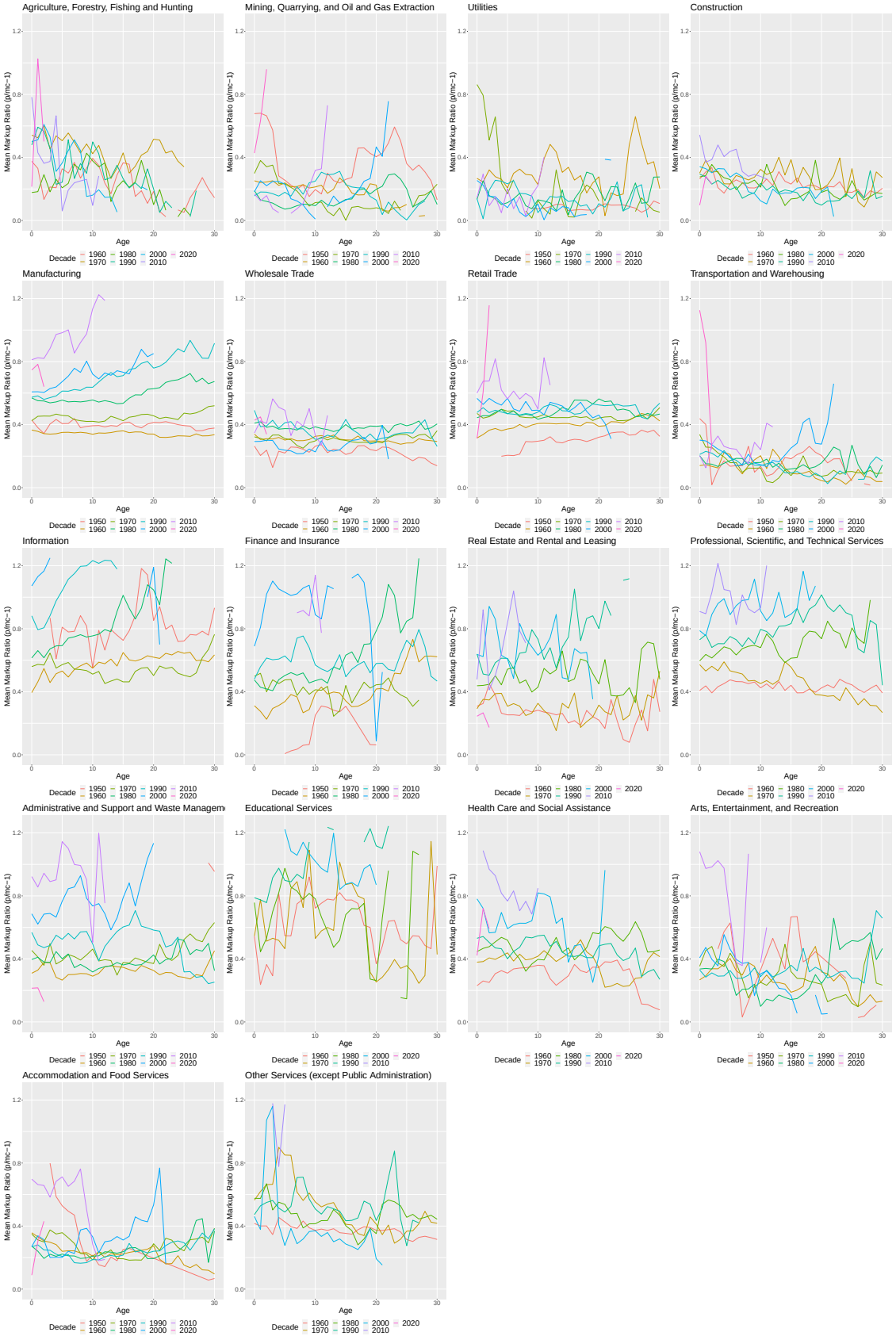


Figure 20: Firms After IPO By Sector and Decade

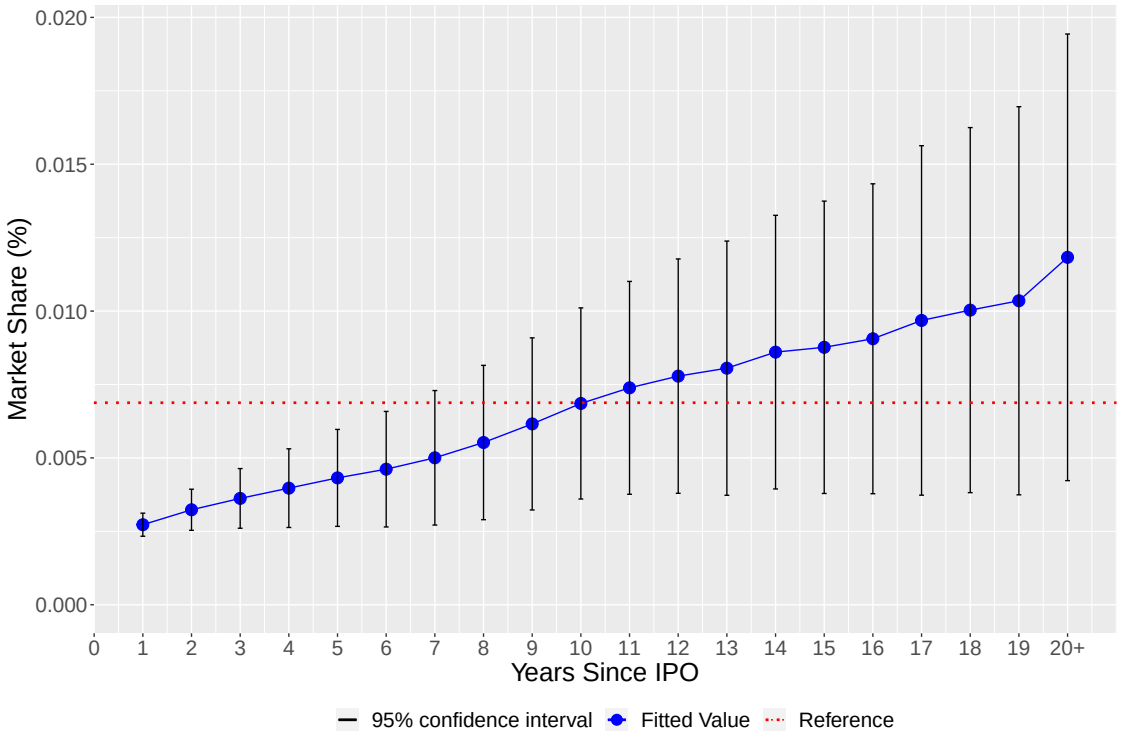


Figure 21: Conditional Average Market Share

G Markup Cyclically: Additional Figures

Figure 22: Markup Response to Supply Shocks by Sector

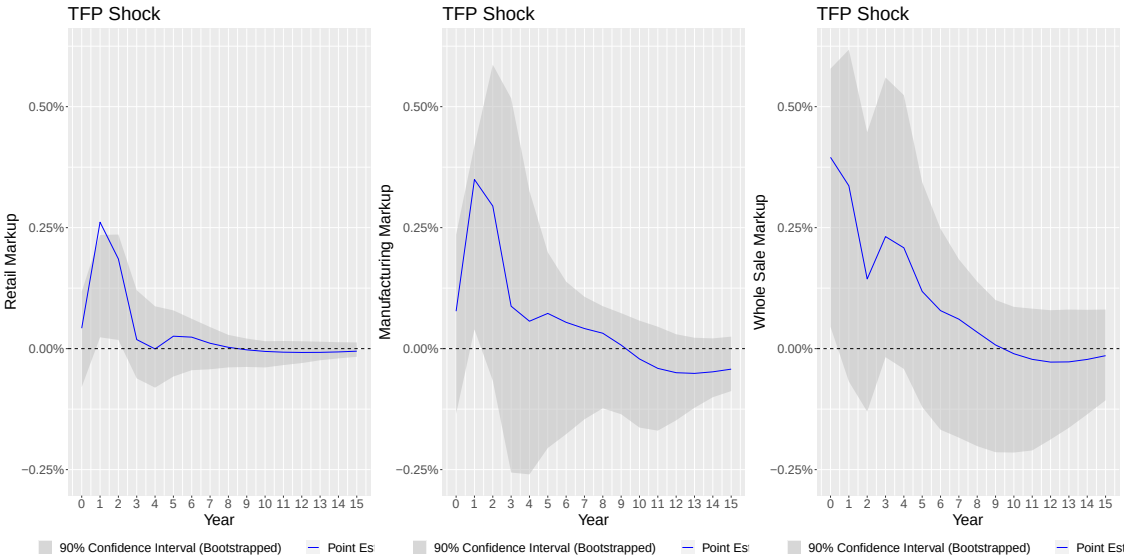


Figure 23: Markup Response to Demand Shocks by Sector

