Macroeconomics and Market Power: Facts, Potential Explanations and Open Questions

Chad Syverson

University of Chicago Booth School of Business and NBER

Economic Studies
at BROOKINGS

The Brookings Economic Studies program analyzes current and emerging economic issues facing the United States and the world, focusing on ideas to achieve broad-based economic growth, a strong labor market, sound fiscal and monetary policy, and economic opportunity and social mobility. The research aims to increase understanding of how the economy works and what can be done to make it work better.

Contents

STATEMENT OF INDEPENDENCE

The past couple of years have seen renewed interest in the potential macroeconomic effects of market power. Empirical investigations have found broad growth in measured profit rates, price-cost margins, and concentration since at least 2000 if not earlier. Those upward shifts were accompanied by drops in measured investment rates, firm entry rates, and labor's share of income. These patterns have sparked robust debate about whether the influence of monopoly power has grown beyond its traditionally studied microeconomic realm of the single industry or market and into the economy overall. If average levels of market power have indeed grown across the board, this is likely to degrade key metrics of economy-wide wellbeing including investment, innovation, total output, and the distribution of income.[1]

Examples of research accompanying this heighted concern include Furman and Orszag (2015), who document a general increase in concentration and note concomitant upward trends in profits in earnings inequality; De Loecker and Eeckhout (2018) and De Loecker, Eeckhout, and Unger (2018), who find that price-cost margins have been on the rise for decades; Gutierrez and Philippon (2017a and 2017b), who find a relationship between market concentration and reduced capital investment; Barkai (2017), who shows that sales concentration within an industry is tied to a reduction in the share of industries' income paid to labor; and Eggertsson, Robbins, and Getz Wold (2018), who demonstrate that a neoclassical model augmented with market power and a declining natural rate of interest can quantitatively mimic observed trends in markups, asset prices, and factor income. This research has accompanied heightened discussion of the issue in the popular press and in policy circles. A few examples among many are *Economist* (2016); Jarsulic, Gurwitz, Bahn, and Green (2016); and Baker (2017).

Economists have studied the causes and consequences of market power throughout the history of the discipline. The recent surge in interest is distinctive in its macroeconomic focus. Prior market power research has largely been the domain of microeconomists who trained their analytical microscopes on individual industries or markets.[2] Decades of microeconomic study have built a knowledge base, formed modeling conventions, and standardized empirical practices. For various reasons, the recent macro-oriented work has often departed somewhat from these established practices. Part of this is surely tied to the obvious difference in the scope of analysis; things that can be done relatively straightforwardly for an individual market are not so easy to do at the economy-wide level. But there are other differences too.

This essay is an attempt to assess macro market power research from the perspective of someone who has worked primarily in microeconomic frameworks when studying market power. At the same time I may be more familiar with the macro literature than many who have worked in the micro market power literature (due primarily to macro-oriented research albeit not dealing with market power per se). I hope in the process to pull the two bodies of work somewhat closer together.

. . .

1. A related debate, though one less participated in by economists, involves the consequences of potential broad-based concentration of not just economic but also political and cultural power (e.g., Khan (2017)).

2. A notable exception is Harberger (1954), who performed a calculation of the aggregate deadweight loss due to market power. He found small effects: "no more than a thirteenth of a percent of the national income" (p. 85). It is plausible that this result had a part in making market power a low priority for macroeconomists until recently.

I begin by formally defining market power. This serves as an empirical touchstone as I next discuss the measurement of market power in the recent literature, paying special attention to two issues raised by this work: the use of concentration as a measure of competition and the quantitative implications of estimated increases in markups. After this more general methodological discussion, I overview some of the more recent and stimulating examples of macro market power research, as well as critiques of this work. As part of this, I characterize the congruencies and incongruencies between macro evidence and micro views of market power and, when they do not perfectly overlap, explain the open questions that need to be answered to make the connection complete.[3]

To preview my conclusion, I believe the macro market literature has very helpfully established and collected an important and provocative set of facts, some developed itself and some built closely upon previous work. The literature has also done a service by drawing plausible connections among these facts and showing how they might be tied to increases in the average level of market power. However, I believe the case for large and general increases in market power is not yet dispositive. There are empirical holes to be filled and plausible alternative stories (some with evidence of their own in their favor) that would first need to be rejected.

To be clear, this is not to say that I believe the case for market power has failed or ought to be rejected. It remains a leading candidate explanation for several trends in the data, and there are empirical results that support it. Certainly more work along these lines is warranted. Rather, I do not think there is yet a rich enough collage of evidence to justify presuming aggregate market power has risen substantially and is responsible for the aforementioned trends. To the extent that there is a case for aggregate market power effects, moreover, the magnitude of these effects is still far from established. It spans possibilities from a trivial blip to massive.

I see the ambiguity of the literature as partly a natural result of the sheer novelty and thinness of this line of work, partly inherent measurement difficulties, and partly a divide between the macroeconomic approach toward empirical work on the subject and the practices honed in microeconomic studies. Given this ambiguity, I wholeheartedly support additional research on the subject that both expands the breadth and depth of specific questions being asked. I also encourage the macro market literature to borrow from the practices of the micro literature when practical and warranted. In the meantime, there is a case for caution before calling for policy changes (antitrust or otherwise) that assume as fact a sizeable, across-the-board increase in market power.

I want to be clear, however, that this is my own interpretation of the evidence. Others may draw different conclusions and policy recommendations.

. . .

3.   For obvious reasons, I focus in this essay on research that involves broad-scoped empirical examinations of market power and its effects. There has been and continues to be lots of work examining market power in specific industries or markets. While some have argued for broader implications to be drawn from these market-specific studies—see for example the reaction of Scott Morton and Hovenkamp (2018) to studies like Azar, Schmalz, and Tecu (2018)—there does not yet seem to be a consensus about whether market-specific studies themselves, whether alone or in their collective weight, have macroeconomic implications.

I. Defining Market Power

Before discussing specific research, it is useful to be clear about how market power is defined. The textbook definition of market power (literally—see, for example, Pindyck and Rubinfeld (2013) or Goolsbee, Levitt, and Syverson (2016)) is that the firm has the ability to influence the price at which it sells its product(s). In other words, if a firm does not face a perfectly elastic residual demand curve, it has market power. A connotation of this definition, sometimes left implicit, is that the firm uses this ability to hold price above marginal cost.4

This condition defines the existence of market power, but the magnitude of market power is tied to the size of the price-marginal-cost gap at the firm's profit-maximizing output level. This gap's size (typically called the markup when expressed multiplicatively and the margin when expressed as a difference, though there is some variation in usage) depends on the shape of the firm's residual demand curve and summarizes how much market power the firm has. Steeper inverse demand means a larger margin and more market power.

II. Measurements of Market Power in the Macro Literature

In addition to the formal definition of market power, there are many informal definitions of market power and associated metrics. These are often used in popular economic writing but sometimes in economic research as well. Examples include the number of competitors (actual or potential), profit rates, and market entry costs. The most frequently used measure of market power in the recent macroeconomic market power literature is concentration. Each of these alternatives has its merits, but each is also one step removed from actual pricing power. This can lend itself to shortcomings and empirical ambiguities in practice. Because of concentration's centrality in much of the macro market power literature, I will discuss its use as a market power measure in detail.

. . .

4. This is the definition of market power in output markets. Firms can also exercise market power in the markets for their inputs. See, for example, Manning (2003); Azar, Marinescu, and Steinbaum (2017); Benmelech, Bergman, and Kim (2018), and Krueger (2018). I focus here on output market power.

II.A. Concentration as a Measure of Market Power

There are multiple concentration measures, but all summarize the share of market or industry activity accounted for by large firms. The most common are the Herfindahl-Hirschman Index (HHI), which is the sum of firms' squared market shares, and Cn, the combined market share of the largest n firms.

An advantage of concentration as an empirical tool for studying market power is that it is often relatively easy to compute. Markups are a direct measure of firms' abilities to set their prices at some level above marginal cost, and as such a direct measure of market power, but one needs both output and cost information to compute them. Concentration requires only revenues.

On the other hand, concentration is a step further removed from the price-cost margin measure of market power than are other markup alternatives like profit rates. Rather than being directly related to margins, concentration is solely about revenues, and relative revenues at that. For example, a monopolistically competitive market can be very unconcentrated—indeed monopolistic competition is defined by the atomistic nature of firms within it—but still have very inelastic residual demand curves and hence a lot of market power.

Concentration also necessarily requires a market definition, and incorrectly defining a market as too large (small) will cause measured concentration to understate (overstate) true concentration. Shapiro (2018) argues this matters for the macro market power literature in two related ways. First, antitrust practitioners would not consider many of the levels and increases in concentration that have cited by some as evidence of broad-based increases in market power to be particularly high. This is in part because commonly used concentration metrics have been measured in overly broad industry groupings. Second, national concentration measures can be particularly misleading for geographically localized markets. For example, a chain restaurant building stores in a number of local markets would tend to increase measures of concentration computed at the national level even if it reduced concentration in the economically relevant local markets. Rossi-Hansberg, Sarte, and Trachter (2018) find evidence suggesting this is "national concentration, local de-concentration" pattern is occurring in a number of industries.

Perhaps the deepest conceptual problem with concentration as a measure of market power is that it is a market outcome, not a market primitive. Concentration is not an immutable core determinant of how competitive an industry or market is. It is instead a result of that competition. Competition and concentration are certainly related; the nature and intensity of industry competition combines with other supply and demand primitives to determine equilibrium concentration. Competition drives concentration, however, not vice versa.

It is important to recognize that this issue does not simply mean concentration is a noisy barometer of market power. Instead, we cannot even generally know which way the barometer is oriented. Concentration can be associated with less competition, but it can also be associated with more competition.

The former case—the one implicitly relied upon by research that uses concentration as a measure of market power—arises in the Cournot oligopoly model. Consider a set of firms, indexed by i, that produce a homogenous good and compete by choosing quantities. For any given market inverse demand $P(Q)$, where Q is the total quantity produced by all firms, firm i's profit maximization problem will be to choose a quantity q_i to maximize $\pi_i = q_i P(Q) - C(q_i)$, where $C(q_i)$ is a cost function with standard properties and constant or increasing marginal costs $C'(q_i)$. The first order condition implies

$$q_i = \frac{P(Q) - C'(q_i)}{-\frac{\partial P(Q)}{\partial Q} \frac{\partial Q}{\partial q_i}}$$

Note that $\partial Q / \partial q_i = 1$. Dividing both sides by the total quantity gives us an expression for a firm's market share s_i:

$$s_i \equiv \frac{q_i}{Q} = \frac{P(Q) - C'(q_i)}{-\frac{\partial P(Q)}{\partial Q}} \frac{1}{Q}$$

Defining the firm's Lerner index as its price-cost margin as a share of the price, $L_i \equiv \frac{P - c_i}{P}$, and using the definition of the elasticity of demand ε, we have:

$$s_i = L_i \varepsilon$$

Thus firms with higher market shares have higher price-cost margins. Defining the market-wide Lerner index, L, as the share-weighted sum of firms' Lerner indexes, we obtain

$$L \equiv \sum_i s_i L_i = \sum_i s_i \frac{s_i}{\varepsilon} = \frac{1}{\varepsilon} \sum_i s_i^2 = \frac{1}{\varepsilon} HHI$$

Where HHI is the Herfindahl-Hirschman index of concentration.

Thus there is a positive relationship between market concentration and average market power in the general Cournot model. More concentration implies less competition. Furthermore, welfare is lower in more concentrated markets because the deadweight loss associated with price-cost margins.

On the other hand, there is a large class of commonly used industry models that predict a positive relationship between competition and concentration. To make the "concentration is an outcome, not a primitive" conundrum clear, I outline an example here. It is a simplified version of the model in Melitz and Ottaviano (2008).[5]

An industry comprises a continuum of firms of measure N. Each produces a single differentiated variety of the industry good (both firms and varieties are indexed by i, and I is the set of industry firms/varieties). Demand for the industry's varieties is given by the representative industry consumer's preferences:

$$U = y + \alpha \int_{i \in I} q_i \, di - \frac{1}{2}\eta \left(\int_{i \in I} q_i \, di \right)^2 - \frac{1}{2}\gamma \int_{i \in I} (q_i)^2 \, di$$

· · ·

5.   Other examples among many of models in this class include Melitz (2003), Asplund and Nocke (2006), and Foster, Haltiwanger, and Syverson (2008). All involve heterogeneous-cost/quality firms selling differentiated goods.

where y is the quantity of a numeraire good, q_i is the quantity of variety i consumed, and $\alpha > 0, \eta > 0$, and $\gamma \geq 0$.

Utility is thus quadratic in total consumption of the industry's output, minus a term increasing in the variance of consumption across varieties. This last term introduces an incentive to equate consumption levels of different varieties. This makes γ a parameter that summarizes how substitutable varieties are for one another. A larger γ means less substitutability (consumers do not want idiosyncratically large or small quantities of particular varieties). As $\gamma \to 0$, substitutability becomes perfect; only the total quantity of industry varieties determines utility. One could think of γ as a direct preference parameter or, with a bit of conceptual license, as a reduced-form stand-in for other (unmodeled) substitution barriers like trade costs, transport costs, or search costs.6

Firms produce their output variety at a firm-specific constant marginal cost c_i. No company will produce if its cost is too high to make a profit; call the cost at which a firm earns zero profits c_D. Given this and demand as described above, Melitz and Ottaviano (2008) show a firm's revenues and operating profits are given by

$$r(c_i) = \frac{1}{4\gamma}[c_D{}^2 - c_i{}^2]$$

$$\pi(c_i) = \frac{1}{4\gamma}(c_D - c_i)^2$$

Entry into the industry is determined as follows. A pool of ex-ante identical potential entrants decides whether to pay a sunk entry cost f_E to take a cost draw from a uniform distribution $G(c) = c/c_M$, $c \in [0, c_M]$.7 If an entrant chooses to receive a draw c_i, it determines after observing it whether to begin production and earn the corresponding operating profits as above. Because only potential entrants receiving draws $c_i \leq c_D$ will choose to produce in equilibrium, the expected value of paying for a cost draw is the expected operating profit conditional on $c_i < c_D$. This yields a free-entry condition:

$$\int_0^{c_D} \frac{1}{4\gamma}(c_D - c)^2 \frac{1}{c_M} = f_E$$

This condition pins down the equilibrium zero-profit cost draw. Solving gives

$$c_D = (12\gamma c_M f_E)^{\frac{1}{3}}$$

Note that the maximum cost at which a firm can profitably operate, c_D, moves in the same direction as γ. That means if substitutability increases (γ falls)—consumers are more willing or able to shift to different firms/varieties—it becomes harder for higher-cost firms to operate.

Melitz and Ottaviano show the average price-cost margin is

. . .

6. Melitz (2003) explicitly models trade barriers. Goldmanis, Hortaçsu, Syverson, and Emre (2010) microfound variety differentiation through search costs. Decreases in these substitutability barriers yield the same comparative statics as this model (an increase in substitutability reduces markups but increases competition).

7. Melitz and Ottaviano have a more general Pareto distribution, $G(c) = (c/c_M)^k$. I exposit the uniform distribution case, where k = 1, for simplicity.

$$\bar{p} - \bar{c} = \frac{3}{4} c_D - \frac{1}{2} c_D = \frac{1}{4} (12\gamma c_M f_E)^{\frac{1}{3}}$$

The margin falls as varieties become more substitutable. In the perfect-substitutability limit, the margin is zero. Thus increased substitutability, which can be thought of as an increase in competition, reduces margins.

From above, the derivative of a firm's revenues with respect to its cost is

$$\frac{dr(c_i)}{dc_i} = -\frac{1}{2\gamma} c_i$$

Not surprisingly, higher-cost firms have lower revenues and market shares, but they lose sales and market share at a higher rate when γ is smaller. The lowest-cost firms account for a greater share of industry sales when substitutability is high.

We now have two results from this model regarding an increase in substitutability: it reduces price-cost margins, and it increases concentration. Therefore in contrast to the Cournot case, the model predicts a negative correlation between market power and concentration.

Two other predictions of the model are relevant to this discussion. One regards the average profit of firms that operate in the market. This is given by

$$\bar{\pi} = \left(\frac{c_M^2 f_E^2}{12\gamma} \right)^{\frac{1}{3}}$$

As substitutability/competition increases (γ falls), firm profits actually *increase*. More intense competition reduces the range of operating cost draws that are profitable, reducing successful entry rates. The free entry condition requires that the entry cost equal the product of two values: profits conditional on having a cost draw low enough to operate, and the probability of receiving a low enough cost draw. This implies that as increased competition reduces the probability of an entrant receiving a low enough cost draw, profits conditional on operating must rise. Interestingly, then, higher profits among firms in the market are not a sign of less competition, but more. (Profits rise despite the lower margins because quantities sold increase markedly as substitutability/competition rises.)

The second additional prediction ties these changes in market outcomes to welfare. Plugging equilibrium quantities produced into the utility function shows that welfare is

$$U = 1 + \frac{1}{2\eta} (\alpha - c_D) \left(\alpha - \frac{2}{3} c_D \right)$$

which is negatively related to c_D. Because c_D falls when γ falls, welfare rises with decreases in substitutability/competition γ. This is not surprising, given that more competition reduces margins and therefore the deadweight loss associated with them. An important contrast is that here, the relationship between concentration and welfare is diametrically opposed to that from the Cournot model.

The negative relationship between market power and concentration is not just a theoretical curiosity. Many empirical studies in varied settings have found that greater substitutability/competition (from, say, reductions in trade, transport, or search costs) shift activity away from smaller, higher-cost producers and toward larger, lower-cost producers. I have

participated in that literature. For example, in Syverson (2004a, 2004b) I show that increases in the ease with which consumers can substitute among producers—spatial differentiation is limited, or products are more physically similar—force out the least efficient producers and increase skewness in the size distribution. In Goldmanis et al. (2010), we demonstrate that search cost reductions reallocate market share toward lower-cost and larger sellers, increasing market concentration even as margins fall. These papers are just a handful of a sizable body of work; it is not an exaggeration to say that there are scores, perhaps hundreds, of such studies. Some focus on specific industries, others more broadly.[8]

Perhaps most relevant to the current discussion are Autor et al. (2017) and Crouzet and Eberly (2018). These studies find (the latter for some sectors, at least) patterns of simultaneous concentration and productivity growth in settings that speak very directly to the recent macro market power literature as well. I will discuss each in more detail below.

A note on the split in empirical practice between the macro and micro market power literatures is warranted here. The logic of misleading concentration is a big reason why the field of industrial organization essentially stopped comparing market outcomes like prices, margins, and profit rates to concentration levels, especially when making comparisons across markets or industries that differ in demand and technology fundamentals. (This earlier body of research is known in the field as the structure-conduct-performance literature.) Fundamentals drive both the extent of competition and the degree of concentration in the industry, and depending on the circumstances those fundamentals can to lead to either a positive or negative correlation between competition and concentration.

While I would not call for a blanket ban on the practice of using concentration to measure market power, caution about the practice is well warranted. There were good reasons for industrial organization to choose to forgo it (particularly, again, for cross-industry comparisons). Simply put, the relationship between concentration and markups, prices, or profits is a relationship between market outcomes. These can be uninformative or, worse, misleading about the causal effect of competition.

Below I will speak further to what the microeconomic literature typically does to measure market power, whether it is practical for macro-oriented work, and what other alternatives might be available.

. . .

8. Changes in production technologies that increase scale economies can also raise concentration. Unlike increases in product substitutability, which by their nature tend to flatten residual demand curves and therefore reduce market power, scale economies have no direct influence on demand. Thus their equilibrium effect on market power is more ambiguous. However, prices could very well still fall even if markups do not, because the scale economies have reduced marginal costs. Arguably, this mechanism in part accounts for the transformation of the U.S. retail sector over the past several decades, first through the diffusion of warehouse centers and superstores and more recently through e-commerce (Hortaçsu and Syverson, 2015).

II.B. Some Macroeconomic Implications of Estimated Average Markups

As discussed above, markups are the most theoretically direct measure of market power among the empirical outcomes often used to proxy for it. Markup measures figured prominently in one of the most influential set of studies in the macro market power literature, De Loecker and Eeckhout (2018) and De Loecker, Eeckhout, and Unger (2018). Both studies find that measured markups have grown substantially since the early 1980s both in the U.S. and in many other countries around the world. I discuss their particular approach and findings further below, but it is first worth discussing some empirical implications of documented increases in markups that appear to struggle to find support in the changes in markups measured in recent research.

One issue is that the finding of growing markups presents a seeming paradox that I raised in Syverson (2018): there appear to be mutually inconsistent patterns in recent decades in measured inflation, markup growth, and cost growth. I can summarize the paradox using the relationship that price, P, equals markup, μ, times cost, C:

$$P = \mu \cdot C$$

According to firms' profit maximization theory, the cost C ought to equal marginal cost, and the markup μ should be a function of consumers' price sensitivity. However, even if prices are not set to maximize profits, the relationship is still quite general and useful. For any consistently measured price and cost, one can define the markup μ as whatever multiplicative factor makes the relationship hold (μ could even be less than 1 if price is less than cost for some reason). In this sense the relationship is essentially an identity and holds by definition.

In growth rates, the relationship is:

$$Growth\ in\ P \approx Growth\ in\ \mu + Growth\ in\ C$$

The growth rate relationship also holds by definition. It is approximate but close to exact in situations we are interested in, where growth rates are relatively modest.

Consider the empirical patterns observed in each of these growth rates over the past few decades.

The left-hand side, the growth rate of prices, is inflation. Measured inflation has been low over the past decade and a half, both in absolute terms and relative to what past relationships between inflation and forcing variables would have implied.

The first term on the right hand side, the growth rate of markups, has of course been notably high and is the object of focus here.

Finally, there is the growth rate of costs. Two things affect costs: productivity and factor prices. Productivity growth has been in a slump since the mid-2000s. Productivity is inversely related to costs, so when productivity grows more slowly than usual, cost growth is unusually high. As for factor price trends over the past couple decades, wage growth has been slow if anything (more so for the middle and lower end of the distribution than the high end), and interest rates have fallen to historical lows. In isolation, those factor price

patterns would tend to slow the growth rate of cost, but they are countervailed by slowing productivity growth.

We can investigate the net effect of these two influences, at least for labor inputs, by looking at unit labor costs. Unit labor costs conveniently combine both productivity and wage effects on costs. They are the ratio of total compensation per hour worked to labor productivity—the nominal labor compensation required to build one unit of output. According to Bureau of Labor Statistics data, U.S. aggregate unit labor cost growth has stayed relatively steady over the past couple of decades.[9] Nominal unit cost growth for other factors was, if anything, faster than labor unit costs. In fact, it accelerated over the period. The BLS's unit nonlabor payments series (which includes capital payments, taxes on production, and any other nonlabor costs) grew 1.58% per year from 1995 to 2004 and 2.34% per year from 2005 to 2017. Some of this may reflect increased markups showing up in payments to capital, and as such would be accounted for already in the identity above. Nevertheless the patterns make pretty clear that cost growth was steady and perhaps even faster than normal over the past couple decades.

Putting these patterns together, we have unusually low measured price growth in the face of unusually high markup growth and steady (or accelerating) cost growth:

$$Growth\ in\ P \approx Growth\ in\ \mu + Growth\ in\ C$$

[unusually low] = [unusually high] + [steady or accelerating]

This is the paradox. How can two components, one growing unusually fast and the other steadily or even speeding up (markups and costs, respectively), sum to sum to a growth rate that has been unusually low (that of prices)? Something doesn't add up, quite literally.

A potential resolution comes from parsing the types of cost in C. Productivity and unit labor cost measures probably most closely reflect average cost. If prices are in fact typically set in a profit-maximizing fashion, however, they depend instead on marginal cost. If marginal costs were falling more quickly than average costs, it is possible that unit labor cost growth could be steady even as inflation remained unusually low. The former would reflect steady changes in average cost; the latter would reflect faster reductions in marginal cost.

This story has the right qualitative features to resolve the paradox. However, it is unclear that it can quantitatively account for the differential patterns in prices, markups, and costs. A simple decomposition of the price-cost markup, first made by Susanto Basu in a discussion of De Loecker and Eeckhout (2017), is instructive about this.[10]

Rewrite the markup by multiplying and dividing it by average costs:

$$\mu \equiv \frac{P}{MC} = \frac{P}{AC}\frac{AC}{MC}$$

Multiplying and dividing P/AC by the output quantity makes it clear that this is equivalent to the ratio of revenues to total costs. The AC/MC ratio is, by definition, the scale elasticity

. . .

9. U.S. nonfarm business unit labor costs sectors grew 1.29% per year between 1995 and 2004 and 1.24% per year from 2005 to 2017. Labor productivity growth fell about 1.6% per year between the same two periods, so this accelerating influence on unit labor costs was almost exactly canceled out by slower nominal wage growth after 2004.

10. I am grateful to Susanto Basu for conversations regarding this decomposition.

of the cost function.[11] When marginal costs are less than average costs, average costs are falling in quantity and the scale elasticity is greater than one. If MC > AC, there are diseconomies of scale, and the scale elasticity is less than one.

We therefore have, using ν to denote the scale elasticity:

$$\mu = \frac{R}{TC}\nu$$

Define pure profit's share of revenues as

$$s_\pi \equiv \frac{R - TC}{R}$$

We can then rewrite the markup as:

$$\mu = \frac{1}{1 - s_\pi}\nu$$

Thus the markup must equal the inverse of one minus profits' share of revenue times the scale elasticity. Note that derivation of this expression requires no other assumptions than differentiability of the function that relates output to costs. (This does even not need to be the standard cost function of production theory; i.e., the total cost expression evaluated at the cost-minimizing factor demands.)

This expression reveals an empirical discipline on measures of markups at the firm level. Namely, markup levels must also imply something about profit shares, scale elasticities, or both. If a firm sees a substantial increase in markups over time, there must also be an increase in pure profits' share of its income or in its scale elasticity.

It is often difficult to obtain firm-level estimates of scale elasticities, as typically common technologies must be imposed across firms in order to estimate an elasticity with any precision. Thus investigating the markup-profit-share-scale-elasticity relationship firm-by-firm can be hard. Exploring its aggregate version can still be informative, but this should be accompanied by the caveat that Jensen's inequality implies the average of the firm-level markup-to-scale-elasticity ratios will not equal the analogous ratio computed in aggregate data. I make this aggregate comparison here, noting this proviso.

De Loecker, Eeckhout, and Unger (2018) report that the average firm-level markup in the U.S. grew from 1.21 to 1.61 between 1980 and 2016.[12] Suppose that the production technology remained stable enough over the period so that the scale elasticity didn't change. Plugging this into the relationship above yields a prediction about pure profits' share of revenues in 2016 relative to their 1980 share:

$$\frac{1.61}{1.21} = \frac{\frac{1}{1 - s_{\pi,2015}}}{\frac{1}{1 - s_{\pi,1980}}}$$

. . .

11.   That is, the inverse of the elasticity of costs with respect to quantity. This is easy to verify. The elasticity of costs with respect to quantity for any differentiable cost function $C(Q)$ is $C'(Q)(Q/C) = MC(1/AC)$. For homothetic production functions, the scale elasticity equals the returns to scale of the production function.

12.   De Loecker, Eeckhout, and Unger (2018) report several sets of markup estimates. I am using their benchmark "PF1" specification. The alternative estimates exhibit similar quantitative behaviors to those described here.

$$s_{\pi,2015} = 0.25 + 0.75 s_{\pi,1980}$$

Even if profits' revenue share was zero in 1980, the observed change in markups—in absence of any increase in scale economies—would suggest a profit share in 2015 of 25%. Given that this is a revenue share and total aggregate sales (revenues) are roughly double aggregate value added, this implies a 50% profit share of value added in 2015. This is enormous, greater than all capital income combined. Now, as noted there is a nonlinearity wedge between the aggregate profit shares and the firm-level average markups. Unless it is very large, though, the calculation implies a misalignment between markups and profits.

Recall, however, that I assumed in the calculation that the scale elasticity v did not change between 1980 and 2015. It is in fact plausible that scale economies increased over the period. Fixed costs may have grown, or the output product mix may have shifted in composition toward lower marginal cost products (e.g., software, pharmaceuticals). This could bring greater consistency to the relationship between the scale elasticity, profit, and markup changes, as the equation indicates an increase in the scale elasticity coincides with either a greater profit share, a higher markup, or both.

I can do such a calculation using De Loecker, Eeckhout, and Unger's (2018) estimates of the average scale elasticity for firms in their sample (which they found increased from 1.03 to 1.08 during 1980-2016) and Barkai's (2017) measure of pure profit's share (which he found grew from 3% to 16% in value added, and thus from about 1.5% to 8% in revenues, between 1980-2014; I will discuss this study in detail below). Plugging these values into the relationship above and taking their ratio yields:

$$\frac{\mu_{2016}}{\mu_{1980}} = \left(\frac{1 - s_{\pi,1980}}{1 - s_{\pi,2016}}\right)\frac{v_{2016}}{v_{1980}}$$

$$\frac{1.61}{1.21} = \left(\frac{1 - 0.08}{1 - 0.015}\right)\frac{1.08}{1.03}$$

$$1.33 = (1.08)1.05$$

$$1.33 = 1.14$$

While the relationship is still some distance from implying consistency, it is closer to equality, suggesting that growth in scale economies are part of the story. In addition, there is the caveat that I am mixing aggregates and firm-level averages when the relationship should hold firm-by-firm.

The markup-profit-share-scale-elasticity relationship is a tool that can be applied more generally, whether it be across firms or over time. The necessary inputs are generally feasible to obtain in the data. Pure profit rates are measurable, and one can use production function estimation techniques to obtain scale elasticities. There are practical hurdles; measuring profits require assumptions about how to measure capital's competitive return, and scale elasticities must often be estimated for pooled sets of producers rather than firm-by-firm. Nevertheless, the relationship imposes a useful consistency check on empirical estimates in this area.

III. Reviewing the Macro Market Power Literature

In this section I sketch out some of the main results of the macro market power literature and the methods used to obtain them. These papers establish or assemble a set of facts regarding aggregate trends in markups, market structure, factor income shares, and industry dynamism over the past 20 years or more. They present the case that some or all of these changes reflect broad increases in market power throughout the economy. Collectively, they lay out a compelling body of evidence to make such connections plausible.

As with almost all empirical work, ambiguities and measurement problems remain. In fact, some have suggested drawing a considerably different inference from the documented empirical patterns. I discuss elements of these critiques and alternative interpretations below as well.

I am not in any way attempting a comprehensive review of work on the topic here. Rather, I want to concisely convey some of the literature's more influential empirical approaches, results, and critiques in order to build a discussion around them.

De Loecker and Eeckhout (2018) and De Loecker, Eeckhout, and Unger (2018). These papers estimate price-marginal cost markups in the U.S. (De Loecker, Eeckhout, and Unger 2018) and in multiple countries around the world (De Loecker and Eeckhout 2018). They use what researchers typically refer to as "accounting data"; essentially, these are data from firms' financial statements. The U.S.-centric study uses the Compustat database, comprising the harmonized financial reports of publicly listed companies for the past several decades. The world study uses Thomson Reuters Worldscope, which spans over 100 countries and contains income statements for publicly traded companies mostly, but it does also include some private firms.

The simplest method with which one might use accounting data to measure markups is to construct an accounting-based proxy for the firm's price-marginal-cost markup. A logical proxy is the ratio of revenues to total variable costs, the equivalent to the ratio of price to average variable cost. Average variable cost does not of course generally equal marginal cost, but marginal cost is very hard to measure directly. Only when marginal cost is constant at all quantity levels are they equal. They diverge to the extent that "inframarginal marginal cost" differs from marginal cost at the firm's observed output. Aside from this, the proxy approach also assumes variable costs are measured without error. This is a high hurdle, as accounting cost categorizations do not make it easy to consistently separate variable from fixed costs.

De Loecker, Eeckhout, and Unger move beyond the simple proxy approach to obtain their estimate of markups. They use a firm-level variant of a method Robert Hall developed and applied to industry-level data (Hall (1988) and Hall (2018)).[13] It is more sophisticated than
. . .

---

13. Hall (2018) uses industry data and finds mixed support for increasing markups. He estimates an average trend in measured markups between 1988 and 2015 that is positive but statistically insignificant (0.6% annual growth, standard error = 0.5%). Multiple measures of returns to capital rise. There is a low correlation between the levels and growth rates of three measures of market power he constructs, but there is a modest positive correlation between concentration and measured markups in his sample.

the proxy approach but shares some elements. Hall (1988) shows that under cost minimization, for any variable input (an input that is freely adjustable by firms within any given period, as opposed to quasi-fixed inputs as many forms of capital are often thought to be), the firm's markup will equal the ratio of two values: the elasticity of output to that variable input, and the share of revenues the input is paid. That is,

$$\mu = \frac{\beta_v}{s_v}$$

Where μ is the (multiplicative) markup, β_v is the elasticity of output with respect to the variable input v (from the firm's production function), and s_v is the share of revenues paid to the variable input supplier.

De Loecker, Eeckhout, and Unger assume that firms' cost of goods sold (COGS) measures their variable input use. They estimate a production function by regressing revenues on COGS and the book value of capital for all firms in an industry. This yields an estimate of the elasticity of output with respect to COGS.14 The other piece of information necessary to estimate the markup, the share of revenue paid to COGS, is observed directly in the data. They take the quotient of these two elements to obtain an estimate of the markup for every firm-year in their data. (The elasticity β_v is restricted to be the same across all firms in an industry or industry-year depending on the specification. The revenue share s_v is firm-year specific.)

An attention-getting headline number from De Loecker, Eeckhout, and Unger (2018) is that the revenue-weighted average markup in the U.S. climbed from about 1.2 in 1980 to 1.6 in 2014. This is a sizeable increase in markups with potentially broad effects. They also find increasing skewness in the across-firm distribution of markups over that period, with average markup growth coming from a spreading of the right tail and a shift in revenue shares toward higher-markup firms. Indeed, the median firm-level markup remained essentially constant throughout the time period.

The broader, world-based study De Loecker and Eeckhout (2018) finds a similar-sized increase in the size-weighted markup, rising from 1.1 in 1980 to 1.6 in 2016. Some systematic variations in this trend exist across continents, however. While Europe, North America, Asia, and Oceania saw rather steady increases over 1980-2015, average markups in South America had little discernable trend. African markups jumped up between 2000 and 2005, but were level before and after.

One of the most compelling elements of these studies is their use of a measure of price-cost margins to gauge market power. As noted above, this is the most direct measure/test of market power among the empirical outcomes often used to proxy for it. Margins are the most theoretically direct measure of the existence and size magnitude of market power. Furthermore, the studies show this growth in estimated market power accompanied nontrivial shifts in important aggregates like higher profit rates and reductions in labor's share of income.

. . .

14.   Production function estimation is itself the subject of a large methodological literature and raises additional measurement
        issues beyond the scope of our discussion here.

In terms of vulnerabilities, the studies use of accounting data has raised some critiques. Accounting data are not constructed for the sake of measuring economic markups, and the reality is that imperfections in measurement are going to exist any time they are used. A prominent issue of debate is the use of COGS to capture variable input use. There are two primary categories of costs reported in accounting data, COGS and SG&A (selling, general, and administrative expenses). COGS includes direct costs associated with purchasing and transforming inputs into the product a company sells, and as such are thought to mostly be composed of variable costs. SG&A includes most other costs, and as such capture many fixed costs. That said, there are several categories of costs that companies might report in SG&A that plausibly scale with the size of operations, and as such are variable costs. Similarly, some costs in COGS might arguably be fixed. Also problematic is the fact that accounting standards actually allow classification of expenses by COGS and SG&A to vary by sector. In some industries, certain expenses might be recorded as COGS while being classified as SG&A in others. Thus in the end the variable/fixed demarcation is not as clean as one would like it to be for measuring markups.

How one measures variable costs in the De Loecker-Eeckhout approach does matter empirically. Traina (2018) shows that if the sum of COGS and SG&A are used as the variable input measure, both the estimated levels and, more to the point, the changes in U.S. markups fall. Instead of rising from 1.2 to 1.6 over 1980-2015, Traina's alternate markups rise from only around 1 to 1.15. To be clear, this estimate of markup growth could also be flawed because of the imperfect mapping between accounting and economic cost categories. De Loecker, Eeckhout, and Unger (2018) argue as much, and justify their focus on COGS by showing that changes in this input measure are more highly correlated with marginal changes in outputs than are changes in SG&A. In the end, researchers using this approach will be left to make, given the patterns in the data, the best possible choice among imperfect options.

Gutiérrez and Philippon (2017a and 2017b). In a pair of papers, Gutiérrez and Philippon (2017a and 2017b) marshal evidence suggesting market power may be behind the low investment rates (relative to profits and Tobin's Q in particular) observed since the start of the 2000s.

Gutiérrez and Philippon (2017b) run a horse race between alternative hypotheses for low investment relative to Q: financial frictions, changes in the nature of investment (e.g., intangibles replacing measured capital investment or globalization shifting investment abroad), increased short-termism in management, and decreased competition. Each class of explanation has multiple specific measures. They find that, at least in terms of ability to statistically explain the unusually low observed investment rate, intangibles account for about a third of the drop, while corporate ownership structure (what fraction of company stock is held by likely long-term investors) and increased industry concentration explain the rest. Measures of financial frictions have no explanatory power.

Gutiérrez and Philippon (2017a) attack the question of causality more directly, using natural experiments and instrumental variables techniques to link changes in competition to investment. The natural experiments involve two measures of increased competition from Chinese imports. The instrumental variable is a measure of "excess entry" in an industry in the 1990s. The logic of the instrument is that the go-go startup environment of the latter

part of that decade in particular led to a large amount of essentially random volatility in entry rates across markets. They show that the amount of 1990s entry relative to fundamentals (both current and in expectation) is correlated with industry concentration a decade later but uncorrelated with observable shocks that occurred in the interim. Instrumenting for industry concentration using excess entry, they find that concentration is negatively correlated with investment rates.

The investment patterns documented by Gutiérrez and Philippon (2017a and 2017b) vividly demonstrate the potential for market power to not just create inefficiencies and reduce output today but also, through its investment effects, reduce future growth rates. The papers are also persuasive in their case that measured investment is low relative to standard explanatory variables.

A few critiques present themselves here. One is addressed in part in both Gutiérrez and Philippon (GP) papers: the role of intangible capital. If intangible capital has become more important over the past couple of decades, and the composition of investment has shifted toward it as a result, the quantitative response of measured investment (which of course does not include intangibles) to traditional forcing variables would decline. But this would be a measurement change, not necessarily an economic one. GP consider this possibility using proxies for intangible capital (including "tangible intangibles"—the capitalized R&D, software, and artistic originals series constructed by the Bureau of Economic Analysis). They find that intangibles could explain some of the drop in the measured investment rate, but a considerable shortfall remains. Moreover, the shortfall is correlated with increases in industry concentration, implying a connection between the two.

Crouzet and Eberly (2018) examine intangible investment as well. They find intangible investment intensity is highest and grows fastest for the largest and fastest-growing firms in an industry. This demonstrates an important possibility. Intangibles need not just be associated with (or caused by) concentration; they can causally affect industry concentration. Thus intangibles aren't just another factor in addition to concentration that might explain low measured investment. They might actually be affecting concentration directly.

This implies that an intangibles-concentration connection can occur through two mechanisms, with very different economic implications. One is the decreased competition interpretation of GP: increased concentration reduces firms' incentives to invest, and this might be coincidentally (or perhaps even causally) correlated with growth in intangible intensity. The other mechanism acts through productivity gains, reverses the potential causality between intangibles and concentration, and has diametrically opposed implications for welfare. If a company invests in intangibles that allow it to deliver a higher quality product at a lower price (reconfiguring its organizational structure and internal processes, for example), market share will naturally shift toward it, creating coincident growth of intangible intensity and industry concentration. However, this shift would not be reflecting greater market power and reduced economic efficiency. It would instead be efficiency enhancing, as the total resources required to deliver a given amount of product quality (consumer welfare) would have fallen.

Crouzet and Eberly compare sectoral-level trends in labor productivity levels and Hall's (2018) industry-level markup estimates to parse between these alternatives. They find that within the manufacturing and consumer sectors (the latter combining wholesale and retail

trade as well as agriculture), estimated markups were flat over 1990-2015, while labor productivity rose. They therefore attribute the coincident increases in concentration and intangible intensity observed in those sectors to efficiency-enhancing mechanisms. On the other hand, they find that both markups and labor productivity grew steadily in the healthcare and high-tech sectors. This indicates elements of both market power and efficiency gains at work.

These results suggest that GP's findings reflect not an across-the-board influence of rising market power but instead an amalgamation of differing mechanisms with quite different economic interpretations. This suggests that some recognition of the heterogeneity underlying aggregate patterns is warranted, both for understanding the phenomenon and drawing welfare implications. It also vividly evokes the aforementioned issues involved with concentration as a measure of market power. That said, to have macro effects market power need not be rising in every sector of the economy. It simply needs to be rising in a sufficiently sized subset sectors so that any countervailing changes in other sectors do not cancel this increase out.

Furman and Orszag (2015); Barkai (2017); and Eggertsson, Robbins, and Getz Wold (2018). Furman and Orszag (2015) notes multiple coincident trends in macro aggregates: increased earnings inequality, increased skewness in returns on capital among publicly listed firms, and declines in labor mobility and business dynamism (entry, exit, and job turnover). The authors argue that a concise explanation for these disparate patterns is an increase in rents accruing to companies. They cite increased industry concentration and an associated gain in market power as the likely source of these rents. They attribute the wage and earnings patterns to successful companies sharing some of these rents with their employees.

Barkai (2017), like Furman and Orszag (2015), pays particular attention to market power's potential effects on the distribution of income. It decomposes aggregate factor income into three elements: labor's share, capital's share, and pure profits. Labor income is taken directly from national income accounts and is therefore measured in standard ways. To compute capital income, Barkai multiplies the observed aggregate capital stock by a user cost of capital. The user cost equals a real interest rate (constructed as average blue-chip bond yields in a period minus a measure of expected inflation) plus a measure of the depreciation rates. This user cost is supposed to reflect the competitive return earned by capital inputs. Any remaining income is considered pure profit.

The results indicate that the past three decades' drop in the share of income paid to labor was not accompanied by an equal-sized increase in capital income. Indeed, capital's share also dropped, albeit only slightly. This means the pure profit residual must have increased, and it did so substantially. Pure profits grew from 3% of national income in 1985 to 16% by 2014. The study ties this shift in factor income shares to market power using regressions conducted at the 6-digit NAICS level. He finds that industries that saw larger increases in concentration saw bigger drops in labor's share of income. He interprets this as evidence that declining competition has been responsible at least in part for the secular decline in labor's share.

Eggertsson, Robbins, and Getz Wold (2018) argues that the standard neoclassical model, augmented with increasing market power/markups and a decreasing natural rate of interest, can qualitatively and quantitatively explain a number of empirical phenomena: the falling labor's share of income, the increase in the pure profit rate, growth in the financial wealth-to-output ratio, an increase in Tobin's Q without associated investment, and a divergence between the marginal and the average return on capital. The mechanism, briefly stated, is as follows. Growing market power (what the paper terms the "emergence of a non-zero-rent economy") leads directly to the increase in pure profits through markups. Financial wealth and Tobin's Q both reflect future claims on profits, so these rise as well. The increase in profits' share decreases both labor's and capital's share. Because higher profits increase the return on capital, however, there must be a countervailing influence in order to generate the roughly constant returns that have been observed in the data. This is where the falling natural rate of interest comes in. The paper demonstrates that its model, suitably parameterized, can match the observed trends in markups, asset prices, and factor income.

A strong common thread in these papers is their exploration of the relationship between increased market power and changes in the distribution of factor income, in particular the decline of labor's share and the growth of pure profit. These factor income share trends had also been noted before in separate literatures. While some have raised concerns about specifics of their measurement, the trends have been documented in multiple ways. It is probably fair to say there is a consensus in the profession that both labor's share has been trending down while corporate profits have risen. These papers take those established facts and tie them to increases in market power, with a focus on concentration as a measure of that market power.

Just as with these papers, Autor, Dorn, Katz, Patterson, and Van Reenen (2017) and Karabarbounis and Neiman (2018) relate the decline in labor's share to increases in concentration. Their explanations are different, however.

Karabarbounis and Neiman (2018) explore three possible explanations for a shrinking labor's share and growing "factorless income," the residual income left after removing payments to labor and imputed competitive capital payments. One is the pure profits story of Barkai (2017) and Eggertsson, Robbins, and Wold (2018). The second is the intangible explanation in Crouzet and Eberly (2018), and the third is that the true rental rate of capital has risen. They conclude that both the profits and intangible capital explanations, while consistent with many empirical patterns in recent decades, imply unusual empirical patterns for earlier postwar history. The tight connection to real interest rates of the profits mechanism would have implied profit rates were actually higher in the 1960s and 1970s than today, which seems at odds with direct measures at the time. And the intangible capital explanation would have implied that intangibles were half of the capital stock in the 1960s, something dismissed as prima facie implausible. They note that the mismeasured rental rate explanation avoids these counterfactual predictions and implies some other empirical patterns more consistent with observed data, though they do not have a sharp explanation for what would cause the true rental rate to vary is it would need to in order to produce the macroeconomic outcomes that have occurred over the prior 60-70 years.

Autor et al. (2017) and Bessen (2017) focus on within-industry changes (much as Crouzet and Eberly (2018) do, though for a different mechanism). Their theorized mechanisms involve the interaction of role of fixed costs and heterogeneous (cost/quality/productivity) firms. Higher productivity firms amortize fixed costs over a larger revenue base. In Autor et al., where fixed costs are assumed to be more labor intensive than variable costs, larger firms pay a smaller share of their revenues to labor. (Fixed costs have a common size across firms in their framework.) In Bessen, which relates fixed costs to IT capital intensity, IT-intensive firms are larger, more productive, and have higher operating margins.

These stories raise the same point as the Crouzet-Eberly results about concentration being related to outcomes through mechanisms other than market power. Autor et al. and Bessen argue concentration has grown because changes in market factors have created an environment that increases skewness—in revenues (concentration) and productivity, certainly, and perhaps in other dimensions. Something has flattened firms' residual demand curves or marginal cost curves, be it increased scale economies, network effects, or improved abilities of consumers to find low-cost or high-quality firms. These lead to increased concentration ("superstar firms" in Autor et al.'s parlance) but do not necessarily imply growth in market power. Increased scale economies may come from reductions in marginal cost that reduce the amount of inputs necessary to produce output, an efficiency enhancement. On the other hand, they require enough market power in equilibrium for firms to pay fixed costs and production costs of their inframarginal units. Network effects also have implications for both efficiency and market power. Consumers can obtain a utility benefit from network effects, but network effects can also cause lock-in, which gives firms pricing power. Improving consumers' abilities to choose from whom they buy, if it comes from changes in search, transport, or trade costs, for example, are likely to be efficiency enhancing. The potential for market power gains in such situations is considerably less clear.

Both Autor et al. (2017) and Bessen (2017) present evidence bolstering the case for an efficiency-enhancing mode of concentration being the primary actor in their data. Autor et al. find that industries that saw greater increases in concentration also saw on average faster growth in patent rates, capital intensity, and productivity. Bessen shows IT system use is in fact tied to concentration and more skewed operating margins and productivity levels in an industry.

The Autor, Dorn, Katz, Patterson, and Van Reenen (2017), Bessen (2017), and Crouzet and Eberly (2018) results, to the extent they support that efficiency gains have accompanied increases in concentration, are examples of why, as discussed above, caution is warranted when inferring market power using industry concentration as a metric.

IV. Filling in the Macro Market Power Literature

In my discussion of the macro market power literature above, I described some current vulnerabilities that in my opinion keep the literature's conclusions from being dispositive. In this section I discuss what might be done to fill holes in the literature and round out the evidence in a way that would allow more definitive conclusions.

A logical first place to look for new threads that the literature could pick up is in the best practices of the well-developed microeconomic literature on market power. Unfortunately, it might not be possible to directly replicate the modal practice for measuring market power in microeconomic research.

This approach operates through the recognition that the optimal price-cost markup depends on the slope of the inverse residual demand curve facing the firm. If that slope can be estimated, the implied profit-maximizing price-marginal cost margin can be backed out from that. Most of the literature follows this logic and estimates the demand system for the products in the market (if the products are differentiated, this is typically accommodated by using a discrete choice demand system where the product attributes are included as demand shifters).

While this method of using demand-side data to measure a technological unobservable (marginal costs) may seem unusual, in many settings the richness of the demand system offers the ability to more precisely estimate demand and therefore implied margins in ways that cost data alone could not. Moreover, one can typically jointly estimate both the demand and supply sides by parameterizing costs (again as a function of attributes if products are differentiated) and using the restriction that the observed product price must equal the estimated marginal cost time the profit-maximizing markup implied by estimated residual demand.

All this said, the demand-system-estimation approach may not be feasible in the multiple-industry/market settings in the macro market power literature. Specifying a realistic demand system typically takes a fair amount of knowledge about the nature of the product and the institutional details of the market. It is not practical to do this in studies that look across hundreds of very different markets. Taking an analogous approach with aggregate data would not work either. Pricing power depends on the slope of firms' residual demand curves, not the slope of the market/industry demand curve. Backing out an implied markup and market power from a market/industry demand curve would be conceptually and empirically incorrect.

What is one to do, then? Well, the logic behind using the price-cost margin as the most direct measure of market power is solid. Trying to measure it directly is difficult, though, because marginal costs are very rarely observed even in detailed microdata. Cost data are almost invariably aggregated across units of output, and in any case marginal costs often involve shadow costs that are by nature unobservable.

If microdata is available, one might imagine a more parametric approach to measuring marginal cost whereby a cost function is specified and estimated using observed variation

in costs. Marginal costs would then be obtained by taking the derivative of this estimated function. This approach is limited by several factors, however. Quantity data of some type are required; revenues (which of course confound price and quantity) alone would not suffice. Many producer-level datasets report only revenues. Further, for highly differentiated products there may not be enough data to fully characterize the cost function given the multiple attributes that could shift costs. Both of these reasons make it more difficult to apply this approach as one moves away from commodity-type products and instead toward more differentiated goods. Another issue is that to consistently estimate a cost curve, instruments that exogenously shift quantities are needed. These are not easy to find in many settings.

When it comes to estimating markups for broad swathes of the economy, there may be no silver bullet. One is left to choose from a set of imperfect choices.

The macro market power literature does need to deal with its use of concentration as a measure of market power. The ideal again would be to measure markups accurately. In absence of that, however, several alternatives present themselves. Sometimes one can obtain direct measures of plausibly exogenous differences in competition. In that case, concentration might be instrumented using those measures or, alternatively, those measures could be used directly as explanatory variables themselves. More generally, if concentration is the only possible metric available, researchers should strive to demonstrate using ancillary evidence that increases in concentration do in fact correspond to more market power rather than efficiency in the market(s) they are studying. An example (favoring the efficiency rather than market power story in this case, but the principle is the same) is in Autor et al. (2017). They show that concentration is associated with innovation, capital deepening, and productivity, which bolster the case for efficiency mechanisms. Alternate findings would have supported a market power interpretation.

Another area for ongoing research in the macro market power literature is to more fully characterize heterogeneity, both across and within markets.

Across markets, as the Crouzet and Eberly (2018) results suggest, market power can act broadly within some sectors but not others. They find the healthcare sector, for example, seems to have seen the influence of market power. This is actually supported by market-specific studies in the micro literature, for example Cabral, Geruso, and Mahoney (2018) and those described in Gaynor (2018). On the other hand, their evidence points to the manufacturing and consumer sectors as not showing the signs of market power. Characterizing such differences and explaining where they come from is important for understanding the mechanisms behind, and the effects of, market power in macro settings.

Within industries, the skewness results shown by De Loecker, Eeckhout, and Unger (2018)—that the increase in average measured markups is driven exclusively by increases in the right tail of the distribution—are an example of the necessity of understanding within-industry heterogeneity to grasp what is happening with aggregates. Averages can obscure. Producers in an industry differ markedly in their behavior, including in their responses to common external influences. Market-, industry-, or economy-wide changes do not always, nor likely even usually, reflect a common change experienced by all producers. Rather, they reflect the summation of what are typically very different responses, and this includes reallocations of activity across heterogeneous producers. The experience of the

median producer (or even the average producer, if producers are equally weighted) is not informative about changes at the industry level. One cannot simply rely on producer-level variation "canceling out" when looking at aggregate changes. That variation is what creates the aggregate changes.

V. Conclusion

Where does this leave us? The macro market power literature has offered an immense service by documenting and emphasizing the potential connections between several trends: labor's declining share of income, increasing corporate profits, increasing margins, increasing concentration, slower productivity growth, decreasing firm entry and dynamism, and reduced investment rates. While none of these metrics are perfect, many (but not all) have been replicated in multiple venues with multiple techniques, and as such can be considered quite robust. The fact that the changes in them are so noticeable and have been trending for so long (each for over a decade at a minimum, some approaching four decades now)—often in contrast to very different patterns before—creates an inherent interest and importance to them.

Where the literature, at this point at least, has not yet reached a conclusion is whether and to what extent increases in the average level of market power in the industry is responsible for each or all of these trends. In my opinion, more needs to happen before we can attribute these changes to greater market power. There are still empirical gaps that need to be closed. There are plausible alternative stories, some accompanied by controverting empirical evidence to the market power hypothesis, that need to be rejected.

Again, this is *not* to say that I believe the market power story has been proven wrong. It is still a viable candidate explanation for the documented trends. This might be particularly true for specific industries or sectors. What I do believe, however, is that there is not yet enough supportive evidence to make a broad-based increase in average market power the undisputed leading candidate explanation.

REFERENCES

Asplund, Marcus and Volker Nocke. 2006. "Firm Turnover in Imperfectly Competitive Markets." *Review of Economic Studies*, *73*(2): 295-327.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017. "The Fall of the Labor Share and the Rise of Superstar Firms." NBER Working Paper No. 23396.

Azar, José, Ioana Marinescu, and Marshall I. Steinbaum. 2017. "Labor Market Concentration." NBER Working Paper No. 24147.

Azar, José, Martin C. Schmalz, and Isabel Tecu. 2018. "Anticompetitive Effects of Common Ownership." *Journal of Finance*, 73(4): 1513-1565.

Baker, Jonathan B. 2017. *Market Power in the U.S. Economy Today*. Washington Center for Equitable Growth.

Barkai, Simcha. 2017. "Declining Labor and Capital Shares." Working Paper.

Benmelech, Efraim, Nittai Bergman, and Hyunseob Kim. 2018. "Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?" NBER Working Paper No. 24307.

Bessen, James. 2017. "Information Technology and Industry Concentration." Boston University School of Law, Law & Economics Paper No. 17-41.

Cabral, Marika, Michael Geruso, and Neale Mahoney. 2018. "Do Larger Health Insurance Subsidies Benefit Patients or Producers? Evidence from Medicare Advantage." *American Economic Review*. 108(8): 2048-87.

Crouzet, Nicolas and Janice Eberly. 2018. "Understanding Weak Capital Investment: the Role of Market Concentration and Intangibles." Working Paper.

De Loecker, Jan and Jan Eeckhout. 2017. "The Rise of Market Power and the Macroeconomic Implications." NBER Working Paper 23687.

De Loecker, Jan and Jan Eeckhout. 2018. "Global Market Power." NBER Working Paper 24768.

Economist. 2016. "Too Much of a Good Thing." March 26th.

Eggertsson, Gauti, Jacob A. Robbins, and Ella Getz Wold. 2018. "Kaldor and Piketty's Facts: The Rise of Monopoly Power in the United States." Washington Center for Equitable Growth Working Paper.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1): 394-425.

Furman, Jason and Peter Orszag. 2015. "A Firm-Level Perspective on the Role of Rents in the Rise in Inequality." Presentation at "A Just Society" Centennial Event in Honor of Joseph Stiglitz, Columbia University.

Gaynor, Martin. 2018. "Examining the Impact of Health Care Consolidation." Statement before the Committee on Energy and Commerce Oversight and Investigations Subcommittee U.S. House of Representatives. Retrieved October 10, 2018 from https://docs.house.gov/meetings/IF/IF02/20180214/106855/HHRG-115-IF02-Wstate-GaynorM-20180214.pdf.

Goldmanis, Maris, Ali Hortaçsu, Chad Syverson, and Onsel Emre. 2010. "E-commerce and the Market Structure of Retail Industries." *Economic Journal*, 120(545): 651-82.

Goolsbee, Austan, Steven Levitt, and Chad Syverson. 2016. *Microeconomics, 2nd Ed.* Worth Publishers.

Gutiérrez, Germán and Thomas Philippon. 2017a. "Declining Competition and Investment in the U.S." NBER Working Paper 23583.

Gutiérrez, Germán and Thomas Philippon. 2017b. "Investment-less Growth: An Empirical Investigation." *Brookings Papers on Economic Activity, Fall 2017*: 89-182.

Hall, Robert E. 1988. "The Relation between Price and Marginal Cost in U.S. Industry." *Journal of Political Economy*, 96(5): 921-47.

Hall, Robert E. 2018. "New Evidence on Market Power, Profit, Concentration, and the Role of Mega-Firms in the US Economy." NBER Working Paper No. 24574.

Harberger, Arnold C. 1954. "Monopoly and Resource Allocation." *American Economic Review Papers and Proceedings*, 44(2): 77-87.

Hortaçsu, Ali and Chad Syverson. 2015. "The Ongoing Evolution of US Retail: A Format Tug-of-War." *Journal of Economic Perspectives*, 29(4): 89-112.

Jarsulic, Marc, Ethan Gurwitz, Kate Bahn, and Andy Green. 2016. *Reviving Antitrust Why Our Economy Needs a Progressive Competition Policy*. Center for American Progress.

Karabarbounis, Loukas and Brent Neiman. 2018. "Accounting for Factorless Income." NBER Working Paper No. 24404.

Khan, Lina M. 2017. "Amazon's Antitrust Paradox." *Yale Law Journal*, 126(3): 564-907.

Krueger, Alan B. 2018. "Reflections on Dwindling Worker Bargaining Power and Monetary Policy." Luncheon Address at the Jackson Hole Economic Symposium August 24, 2018.

Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, *71*(6): 1695-1725.

Melitz, Marc J. and Gianmarco I. P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies*, 75(1): 295–316.

Pindyck, Robert S. and Daniel L. Rubinfeld. *Microeconomics, 8th Ed*. Pearson.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. 2018. "Diverging Trends in National and Local Concentration." Working Paper.

Scott Morton, Fiona M. and Herbert J. Hovenkamp. 2018. "Horizontal Shareholding and Antitrust Policy." *Faculty Scholarship*. University of Pennsylvania Law School.

Shapiro, Carl. 2018. "Antitrust in a Time of Populism." *International Journal of Industrial Organization*.

Syverson, Chad. 2004a. "Market Structure and Productivity: A Concrete Example." *Journal of Political Economy*, *112*(6): 1181-1222.

Syverson, Chad. 2004b. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics*, 86(2): 534-550.

Syverson, Chad. 2018. "Changing Market Structure and Implications for Monetary Policy." Remarks at 2018 Jackson Hole Economic Policy Symposium.

Traina, James. 2018. "Is Aggregate Market Power Increasing? Production Trends Using Financial Statements" Stigler Center New Working Paper Series, No. 17.

B | Economic Studies
at BROOKINGS

The Brookings Economic Studies program ana-
lyzes current and emerging economic issues fac-
ing the United States and the world, focusing on
ideas to achieve broad-based economic growth, a
strong labor market, sound fiscal and monetary
policy, and economic opportunity and social mo-
bility. The research aims to increase understand-
ing of how the economy works and what can be
done to make it work better.

Questions about the research? Email communications@brookings.edu.
Be sure to include the title of this paper in your inquiry.