From Good to Bad Concentration? U.S. industries over the past 30 years
by Covarrubias, Gutierrez, and Philippon

Discussion by Janice Eberly, Kellogg School of Management,
Northwestern University and NBER
NBER Macroeconomics Annual 2019

The paper by Covarrubias, Gutierrez, and Philippon provides many useful insights into the rapidly emerging literature on rising concentration in US industries. Importantly, it catalogues some important empirical shortcomings in the literature. It also provides clarity on conceptual issues that have created confusion. The paper goes on to make two types of original empirical contributions. In the first, it focuses on two categories of explanations for rising concentration: "good" and "bad". The former is associated with technological change that increases the elasticity of substitution among goods ($\sigma$, and hence greater competition) or increases firms' accumulation of intangible capital ($\gamma$, perhaps associated with network externalities and returns to scale). "Bad" concentration is instead associated with rising barriers to entry, $\kappa$. The authors argue that the data tend to favor good concentration earlier in their data sample - through 2000. Thereafter, there is increasing evidence of barriers to competition. The collection of evidence, while not dispositive, moves the weight of the evidence toward market power explanations, especially later in the sample. The last part of the paper takes a different tack. Instead of looking for indicators of market power to explain a broad range of facts, the paper looks at combinations of explanations by industry, and argues that there is merit in several of them, and the results vary by industry. I will argue that these last insights are especially helpful, as the macroeconomic data are unlikely to be captured by a single simple narrative.

1 Rising concentration

The paper first documents the fact that has captured the imagination of many authors: concentration seems to have risen in many U.S. industries. This tendency has been found using various measures, in different data sets, at varying levels of aggregation. Using Compustat data, for example, I find (Crouzet and Eberly, 2019) that the Herfindahl index of sales in Compustat firms has increased by at least 50% in 75% of US industries since mid-1990’s. Similar results are cited in the paper. What to make of that finding is much less clear, and the explanation for rising concentration is important. Some explanations, such as rising market power, may give rise to economic inefficiencies (though they need not), while others, such as the rise of more productive "superstar" firms, may be efficient. Observing concentration alone is not sufficient to know.

A growing body of research has developed to understand the reasons behind the rise in concentration. This paper first does a service to this literature by codifying methodological points. The first is on measuring markups, which are often used as a way of measuring market power. The simple point is that fixed costs make it difficult to measure markups. If the researcher wants to measure the markup over variable cost, the presence of fixed costs can
confound the estimate, especially in accounting data where fixed and variable costs are not well-differentiated. In practice, using Compustat accounting data for publicly traded firms, a simple measure of markups is the retail value of output (Sales) over the cost-of-goods-sold (CoGS), treating CoGS as the measure of variable cost. The other component of costs in Compustat is Sales, General, and Administrative Expense (SG&A). Some researchers treat this as a fixed cost, and hence excluded from the denominator of the markup calculation. Using that approach, the estimated markup is large and rising, as the components of SG&A have risen over time, reflecting the growing role of intangible capital, marketing expenses, and other expenses. CGP argue that these expenses are entirely variable and instead add them to the denominator of the markup. Using this approach, they find a smaller, though still rising, markup. In practice, the truth is likely in between, as the SG&A category includes some expenses, such as human resources, branding, and marketing, that have some element of overhead or fixed costs. But the discussion in the paper frames the issue and scales it quantitatively. The ultimate answer to the data quandary probably lies in getting better measures of market power than allowed by currently available accounting data.

The second methodological improvement of the paper is moving beyond macro data and using industries to measure concentration. This is the approach used in (Crouzet and Eberly, 2019), recognizing that firms’ markets (and hence market power) are likely defined at a disaggregated level. I expect that economists and lawyers who focus on antitrust issues would find the SIC and NAICS industry definitions to be frustratingly crude and only tangentially related to the true nature of competition between firms. However, if these industry classifications were arbitrary, then we’d expect to find little when examining concentration and the characteristics of these industries. To the contrary, the authors find much. Nonetheless, it is worth remembering that firms compete across industries - for example, Google and Amazon are increasingly competing for advertising dollars, but they are in quite different industries based on the SIC and NAICS definitions (internet search versus electronics retailing).

2 Why has concentration risen? (1) Looking at individual explanations

The first part of the empirical section focuses on the potential explanations for the rise in concentration. These explanations are mapped into parameters of the model laid out in the paper: the demand elasticity and the returns to scale ($\sigma$ and $\gamma$ that lead to ”good concentration” from substitutability and intangibles), and barriers to entry ($\kappa$ that leads to ”Bad concentration,” from barriers to entry). While the paper has a useful model that can give rise to concentration, the mapping to the data is less clear in two dimensions. First, as the authors note, the explanations are not mutually exclusive, so they could be operating in concert and even interacting. Intangibles, like software and intellectual property, are not simply substitutes with physical capital in the production function, and may have properties, such as scalability and excludability, that are more connected to increasing returns and network externalities, for example. Second, the empirical tests generally focus on objects like market leaders and churn that may not connect directly to a single parameter or property of the model. Hence, it can be hard to interpret the empirical findings as evidence for a single
driver of concentration, or even a category such as "good" or "bad". They are more generally associated with market power, and hence indicative of a root cause.

2.1 measurement questions

The paper implements a battery of tests to examine the relationship between concentration, markups, and various metrics, such as market share, entry/churn, and investment. Measurement is always a concern, but I would note two relevant issues in this line of research. The paper reasonably uses investment as an indicator of firm outcomes. However, with intangibles, there are a number of potential measures at the firm, industry and aggregate level. Aggregation can be problematic empirically, as these measures have different trends: intangibles are generally growing, while physical capital has tended to stagnate or even decline. Hence, for example, the choice of numerator and denominator in investment rates requires some care, and it is not hard to induce mechanical trends by using different measures of investment. In particular, the capitalized spending on SG&A, as in (Peters and Taylor, 2017), has no counterpart in the industry and aggregate data, and its trends can be very large. These choices can have consequences for how we view industry leaders, for example, as they tend to be intangibles intensive.

The analysis of market leaders is one of the most novel interesting parts of the empirical analysis. Once concern, though, is the substantial evidence (Stulz and Kahle, 2017) documenting the shrinking number of public firms. Since the Compustat data include only publicly traded firms, the denominator is shrinking in all of the concentration measures. This decline is also relevant for the entry and turnover analysis in the paper, suggesting some caution in interpreting these results.

2.2 multiple rather than single explanations

A final question is how to interpret the individual empirical tests examining one hypothesis at a time, even if grouped into "good" versus "bad". There is an Occam’s Razor appeal to finding a single explanation, such as market power, for a panoply of facts. On the other hand, as the authors acknowledge, these explanations are not mutually exclusive, and describe a broad set of inter-related macroeconomic outcomes over decades. For example, intangible capital, such as patents, may create barriers to entry. Similarly, firms may invest in barriers to entry, which are then endogenous. It may be more fruitful to think of the empirical evidence presented here as more of a collage of evidence than single "smoking gun" that favors a particular explanation at a point in time.

As an example, the recent paper by Farhi and Gourio (Farhi and Gourio, 2018) allows for a changing capital share, intangible capital, and a rising risk premium simultaneously in an macro growth model. They find that several forces, especially the rising risk premium, are necessary to explain the macro data. Notably, the rising risk premium needed to match the cost of capital (when the risk free rate is low) already puts downward pressure on capital accumulation. Hence, rents from market power and intangible capital actually work to raise the capital stock to measured levels. This is quite a different mechanism than emphasized in CGP. It may or may not be correct, but it illustrates the potentially complex forces when
more than one explanation is allowed. This is particularly important when the question is quantitative, not just directional.

Modeling multiple explanations can also produce quantitatively important interactions. In current work with Nicolas Crouzet (Crouzet and Eberly 2019), we allow for both market power and intangible capital in a $Q$ model of investment. We specify a revenue function for firm $j$ to allow market power and two types of capital, $K_1$ and $K_2$, which we think of as physical and intangible capital:

$$\Pi_{j,t} = \mu D_t^{\mu-1} (Z_{j,t})^\mu ((1 - \eta) K_{1,j,t}^p + \eta K_{2,j,t}^p)^\mu.$$  

where

- $Z_{j,t}$: firm-level productivity.
- $D_t$: industry-wide demand shifter.
- $\epsilon \geq 1$: demand elasticity; $\mu = \frac{\epsilon}{\epsilon - 1} \geq 1$: markup.

In this setting, we can decompose average $Q_1$, which is empirical average $Q$ for physical capital, into the marginal $q_1$ that determines physical capital investment, and three "wedge" terms. The first two measure the value of intangible capital and the rents from market power, respectively, and the last one measures their interaction.

$$Q_{1,j} = q_{1,j} + q_{2,j} \nu_j + \frac{1}{r-g_j} (\mu - 1) R_{1,j} + \frac{1}{r-g_j} (\mu - 1) \nu_j R_{2,j}$$

Intuitively, the first wedge term captures the value of intangibles that contribute to the value of the firm. The second term calculates the value of the rents from market power. The third term captures the present value of market power rents to intangible capital, and hence the interaction between intangibles and market power. This term can be quite large, since the cost of capital for intangibles tends to be large due to their high depreciation rates. Simple calculations suggest this last term can be the largest wedge, and hence very important for understanding weak investment. Examining either market power or intangibles individually would find a large effect, but miss that there is the important role and interaction with the other.

3 Why has concentration risen? (2) One size does not fit all

The second empirical section of the paper takes a different approach, and rather than looking at individual explanations, looks at many. The authors employ a principal components analysis which, importantly, allows for different loadings by industry. Here the authors identify (ex post) two principal components - one they define as "Good concentration," which intangibles-related, and the second as "Bad concentration," related to barriers-to-entry. The
results demonstrate both multiple factors at play and much heterogeneity by industry. These results are novel and interesting, and a conceptual and empirical pivot from the first part of the paper.

Principal components are notoriously difficult to interpret, so one has to be careful of falling into the "Rorschach test," where we tend to see our priors in the data. That being said, the patterns are interesting. Figures 26 and 27 report the loadings by industry on the first two principal components, described as "Intangibles-related" and "Barriers-to-entry" respectively by the authors. The weightings are reminiscent of the industry patterns in \cite{Alexander2018}, which report the share of aggregate investment accounted for by industry groups over time, reproduced here in Figure 1. The striking fact apparent there is that the share of physical capital has shifted away from production industries, like manufacturing, and toward geographically stationary industries, like utilities and energy extraction, which put in place pipelines and towers. The industries that have seen the most growth in value, such as high tech, have a flat share of physical investment. Instead, these industries have rising intangibles investment. This evidence complements the industry ranking for PC1 in Figure 26, showing the highest loadings in computer and data industries, as well as chemicals (which includes pharmaceuticals - a high patent field). The lowest loadings are on pipelines and utilities, which have more fixed capital. The highest loadings for PC2, which focuses on barriers-to-entry, are in telecom, info-movies, and finance. However, some of the "new economy" industries, such as inf data and inf-publish, as well as chemicals, have high loadings on both PC2 and PC1, as shown in Figure 13 in the paper. These may be industries in which the rise in intangibles generates growth, but also creates excludability, for example by patent protections, which act as barriers-to-entry.

While this analysis is not dispositive, it does point in the direction of looking for complementary explanations for the rise in concentration. The data here are thought provoking and will hopefully give rise to more detailed examination.

4 Conclusions

Covarrubias, Gutierrez, and Philippon have written a conceptually clarifying and helpful paper. They make clear that the rise in concentration on which much research has been focused is consistent with many economic models, which may give rise to efficient or inefficient concentration. Moreover, they provide a guide to pitfalls in measuring markups using accounting data, and while they may not have the ultimate measure at this point, their work suggests caution in interpreting the measures we have. The empirical work tends to move the weight of the evidence toward a role for market power, especially over time. But the principal components analysis makes clear that a single, across-the-economy story is unlikely to capture the richness of the data. However, by providing this lens to focus our view, the results with heterogeneity and complementary hypotheses are actually more convincing, because we more readily see how market power, technological change, and other forces may interact to generate big trends in the macro data.
The distribution of investment across industries, over time

Figure 1: Physical investment is increasingly allocated toward "fixed" industries, like energy and telecomm. High growth industries, like Tech and Health, show little growth in their share of physical capital. Source: Alexander and Eberly (2018).

References


