Increasing Returns to Scale and Markups

Olga Shanks
George Mason University

Abstract

I estimate the aggregate elasticity of scale for the U.S. economy and find that it is 1.1 and has been rising. I arrive at the industry-level and aggregate estimates by applying Olley-Pakes and Ackerberg-Caves-Frazer estimation methods to the data on the U.S. publicly traded companies over the period from 1980 to 2019.

The elasticity of scale in turn serves as an input for calculating industry and aggregate markups. Increasing returns to scale help explain observed increases in markups over the last decades for broad sectors of the economy. My estimate of 1.2 for the aggregate markup is significantly lower than the estimate of 1.6 found in recent literature. The large disparity in markup estimates stems from differences in the treatment of fixed and variable costs and the methodological approach to the calculation of markups.

Motivation

Increasing returns to scale can explain:
- Rising industry concentration
- Decreasing share of labor in total output
- Rising markups

Yet, economists use the assumption of constant returns to scale:
- Autor et al. (2020) use constant returns and require a change in consumer price sensitivity
- Karabarbounis and Neiman (2014) use constant returns and require the capital-labor elasticity of substitution to be greater than one
- De Loecker et al. (2020) argue that markups cause industry concentration
  - Then what causes markups? Is causality reversed?

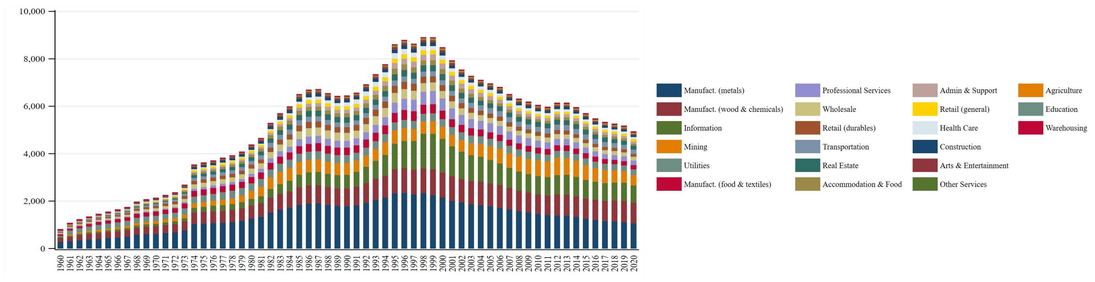
Estimation of Markups:
- De Loecker et al. (2020) argue that the aggregate markup of U.S. firms rose from 1.2 to 1.6 since 1980 to 2016
- Inconsistent with profitability trends
- Treatment of variable and fixed costs
- Long vs. short horizons

Data and Estimation

Data
- Compustat Fundamentals Annual database (U.S. from 1980 to 2019)
- Variable costs: Cost of Goods Sold (COGS) + Selling, General and Administrative (SG&A)
- Capital costs: Property, Plant and Equipment (PPE) × user cost of capital
- Revenues and costs deflated by BEA chain-type price indexes by industry
- Excluded Finance sector (NAICS code 52)
- **5-year rolling periods, e.g. 1980-1984, 1981-1985, etc.**
  - Estimation by industry by year is infeasible due to some industries having too few firms

Estimation of the Elasticity of Scale
- Estimation by industry by 5-year period
- Primary estimation methods:
  - Olley-Pakes (OP)
  - Ackerberg-Caves-Frazer (ACF)
- Invert the investment function to arrive at the productivity shock to eliminate simultaneity bias
- Estimate the probability of exit to eliminate selection bias

Figure 1. Number of firms in Compustat by year by industry.



Results: Elasticity of Scale

- Elasticity of scale is above 1 => Increasing returns to scale
- Divergence after the Internet revolution
- Focus on ACF because it is the most conservative estimate, it allows estimation of standard errors, and it assumes that variable costs are dynamic like capital

Figure 2. Aggregate Elasticity of Scale.

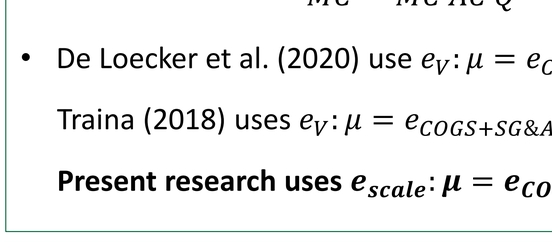
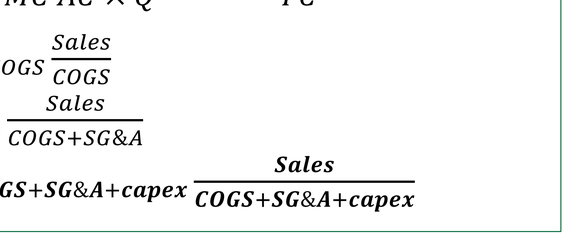


Figure 3. Elasticity of Scale with a 95% Confidence Interval.



Markup Computation

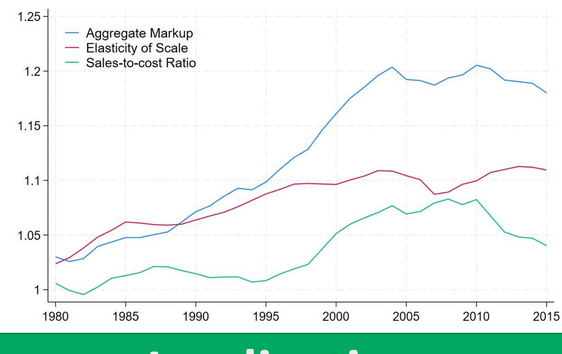
- Cost-minimizing firm:
  $$L(V, K, \lambda) = P^V V + rK - \lambda(Q(\Omega, V, K) - \bar{Q})$$
  Derive from FOCs:
  $$\mu = e_V \frac{PQ}{P^V V},$$
  where μ is markup and e_V is output elasticity of the variable input
- According to Varian (1992), Syverson (2019) and others,
  $$\mu = \frac{P}{MC} = \frac{P}{MC}\frac{AC}{AC}\frac{Q}{Q} = \frac{AC}{MC}\frac{PQ}{AC \times Q} = e_{scale}\frac{PQ}{TC}$$
- De Loecker et al. (2020) use e_V: $\mu = e_{COGS}\frac{Sales}{COGS}$
  Traina (2018) uses e_V: $\mu = e_{COGS+SG\&A}\frac{Sales}{COGS+SG\&A}$
  Present research uses e_{scale}: $\boldsymbol{\mu = e_{COGS+SG\&A+capex}\frac{Sales}{COGS+SG\&A+capex}}$

Choice of Markup Formula

Why include SG&A?
- Firms have the incentive to improve Gross Margin (Revenues - COGS) and they have been shifting costs from COGS to SG&A

Why include capital costs?
- $\mu = e_V \frac{PQ}{P^V V}$ reflects a short-term view, where firms cannot adjust capital
- $e_{scale}\frac{PQ}{TC}$ reflects a long-term view, where all costs can change
- The long-term view is more appropriate for looking at data from 1980 to 2019 and for looking at large firms (most publicly traded firms are large)
- Hardware refresh cycles have been shrinking from 10 to 5 to 3 years
- If capital is omitted, then renting vs. owning real estate results in different classifications of costs for firms in the same industry

Figure 4. Decomposition of the Aggregate Markups.



Implications

Antitrust policies guided by the idea that breaking up large firms should increase and strengthen competition and therefore be efficiency enhancing must face the reality that large size may be an important source of productivity for a firm. Large firms do not get larger only due to the luck of the draw of the total factor productivity, as is often argued, but also because their productivity grows with size.

Contact

Olga Shanks
George Mason University
Email: ostaradu@gmu.edu
Website: www.olgashanks.com
Phone: 571-277-0425

References

1. Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.
2. Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics,* 135(2):645–709.
3. De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics,* 135(2):561–644.
4. Karabarbounis, L. and Neiman, B. (2014). The global decline of the labor share. *The Quarterly Journal of Economics*, 129(1):61–103.
5. Olley, S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications industry. *Econometrica*, 64:1263–1298.