

Empirical Exercise

Coase-Sandor Full-Time Research Professional Application

Summer 2024

1 Introduction

Recent theoretical and empirical research has explored the anti-competitive effects of common ownership – the idea that investors holding shares across multiple competitors in a market can distort incentives for profit maximization. This NBER working paper provides an overview of the underlying theory, the state of the literature, and recent developments.

This exercise uses the framework of this topic to walk through various Stata coding tasks and empirical research questions.

2 Questions & Instructions

1. One of the most often used measures of concentration is the Herfindahl–Hirschman Index (HHI). Using the data we provided, please calculate the HHI for all industries, as specified by their SIC code. Output a table summarizing your results by year.

Hint: The formula for the HHI for industry i is:

$$HHI_i = \sum_{f=1}^N (MS_{fi})^2 = \sum_{f=1}^N \left(\frac{sales_{fi}}{\sum_{f=1}^N sales_{fi}} \right)^2$$

, where MS_{fi} is the market share of firm f in industry i — which is the industry firm f operates in as indicated by its SIC code — and N is the number of firms in that industry.

2. A common criticism of HHI calculation when using sales data from sources such as Compustat is that the total industry output is only measured using public companies, neglecting private companies.¹ In the context of the USA, most companies are private and thus are ignored in such calculations. Therefore, we have provided you with data from the Bureau of Economic Analysis (BEA), which has aggregate industry output data. Please recalculate the HHI using this data.² Compare the results from this calculation to your previous results, using a scatter plot. Do you notice any trends or differences?

Note: The crosswalk between NAICS and SIC codes is not perfect, which means that you might not be able to have a perfect merge. The motivation behind this question is that we would like to see how you would approach "messy" data. The results from this question are not required to answer any of the following questions.

Hint: While there are many different ways to solve this question and no single correct approach, the following might help you get started. We would start this task by focusing on the NAICS codes in the `BEA_gross_outputs.dta` dataset. Then, we would merge the BEA data back to the `sales_profits_data.dta` by year, which can be done with the help of a loop.

¹An interesting discussion of concentration measures and their problems can be found here.

²The data is more aggregated, and thus, do not worry if you can't match every sector.

3. So far, you have calculated the HHI, which measures industry concentration, using the **CommonOwnership.dta** please calculate the Investor Herfindahl–Hirschman Index (IHHI), which measures how concentrated the ownership of a company is. Calculate the average IHHI by industry using the provided SIC codes. Summarize your results in a table by year.

Hint: The formula for the IHHI for firm f is

$$IHHI_f = \sum_{s=1}^N \beta_{fs}^2$$

, where s denotes an investor in that company, and N denotes the set of investors of company f . The variable *beta* in the dataset is the same beta as in the formula above.

4. Now that you have calculated both the IHHI and HHI compare them.³
5. As mentioned in the introduction, common ownership can distort a firms profit-maximizing behavior. As such, we are interested to see if this claim holds true. Similarly, companies without much competition might display different profit patterns, than companies that have lots of competition in their industry. You're free to investigate these claims however you like. How confident are you in your results? You can use all the data at your disposal.
6. If you were to write a paper investigating, the relationship between profits and common ownership, how would you try to approach such a question, what resources would you need and what problems might you face.
7. Do you have any outstanding questions about this exercise? If so, what approaches did you employ to try to answer them and what would be your next steps? If not, please briefly explain your process and the resources you used to complete this task.
8. Are there any Stata coding skills that this coding task did not allow you to demonstrate? If yes, please feel free to mention them and provide some context for us to better understand how and where you used them.

3 Data

Data: In the zip file that you have received in the data folder, you will find four datasets.

1. **CommonOwnership.dta**, is a dataset which provides you with data on S&P500 companies and their investors (this dataset includes only investors that had to fill a 13f with the SEC).
2. **BEA_gross_outputs.dta**, this dataset provides gross output data from the Bureau of Economic Analysis (BEA) for many different sectors.
3. **sales_profits_data.dta**, this dataset contains quarterly sales and profit data for all companies in the **CommonOwnership.dta** data.⁴
4. **2022-NAICS-to-SIC-Crosswalk.xlsx**, this Excel-file provides you with a crosswalk from NAICS to SIC codes.

³We recommend using the sales data we provided to calculate the HHI. However, you can also use the HHI calculated using the BEA data for the comparison.

⁴This data is randomly generate and does not reflect actual sales or profits data for these companies.

4 Deliverables

Please include a writeup describing your approach to the question and presenting your findings. In doing so, please include any relevant tables or figures. There is no minimum or maximum length. There is no single correct way of answering these questions. We are more interested in how you approached the questions than in the actual results. Therefore, if you feel you are stuck with one question or are not confident in your results, please help us understand your process and explain how you approached the question and what difficulties you encountered.

- Please conduct your analysis in Stata, and submit your code and .do files, as well as any outputs, in a .zip file. Please include relevant comments in your .do file so that we can follow your code.⁵
- You may produce your writeup in Word or LaTeX. Please submit your writeup in .pdf format.
- Submit your code and writeup on Box.

⁵In case you do not have access to Stata, you can request a short-term license from Stata Corp.