Case Studies for the Econometrics project work

M1 MBFA

Academic Year 2023/2024

Case 1

A bank wants to incentivise retirement plans. It selects a group of reasonably homogeneous clients for which it has enough data and tests whether follow-up meetings with salesperson affect saving behaviour in general and retirement savings in particular. To evaluate the effectiveness of its policy, the bank calls an economist and hands over the data, the economist randomly assigns the follow-up meeting with salesperson to a subset of clients and waits until more data come in. The economist checks some relevant data, that is the level of income, if the client saved already and how much, if the client already had an active and funded retirement plan, how much the clients earn in income every year. The meeting takes place at time 2, meaning that the economist has two periods prior to the meeting, and she follows the clients for the next three periods. The economist runs through her econometrics class notes looking for an appropriate way to assess the effectiveness of the salesperson meeting, she lists the assumptions that the case at hand checks and starts an empirical analysis. In her experience, she has always found more informative to work with transformed data that abstract from absolute levels, so to really pin down the effect on people's behaviour, rather than picking up economy-wide trends.

Variables list:

- id: client unique identifier
- time: period of observation, from 0 to 4
- female: dummy for gender, female = 1, male = 0
- yincome: yearly income, thousands
- savings: yearly amount of savings, thousands
- retirement: yearly amount in retirement savings, thousands
- meeting: dummy for meeting, 1 if it happened, 0 otherwise

Case 2

The government wants to deploy a green policy that improves existing technologies and reorients research efforts and investments towards more sustainable goods and services. Therefore it enacts a subsidy directed towards those firms that are found to be below the 20% threshold of general R&D investments over profits from last year. Two years after the subsidy was unrolled, Ministries want to know whether the money budgeted for the subsidy was well spent, so they call their chief economist and ask her to look into the matter. She collects enough public and private data on firms from the stock market to run some regressions and investigate the issue. According to the policy, the subsidy goes to green R&D investment, but can also affect other dimensions, like stock performance or general R&D investment. The economist puts together a comprehensive dataset that includes balance sheet and stock market data, from the year prior to the policy to two years after. She obtains series on the number of floating shares, the total market capitalisation, operative costs, total investments in R&D and investments on green R&D, even the length of the ticker for each firm. The economist puts her econometrician hat on and checks how to recover useful information from these limited data, since heterogeneity across sectors, over time, and within firms is so high. In particular, levels are poorly informative, and usually firms take decisions in terms of ratios and shares, so it is important to extract proper information before running the analysis.

Variables list:

- id: firm unique identifier

- time: period of observation
- sectorf: industrial sector
- floating: number of floating shares for the firm
- capit: total market capitalisation, thousands
- opcost: operative costs, millions
- grevenue: gross revenues, millions
- tickerlen: trading ticker length
- invrd: amount of total investment, millions
- invgreenrd:amount of green tech investment, millions

Case 3

An asset manager wants to check what determines the performance of two key financial metrics, so to inform her investment decisions: she wants to look at the determinants of market capitalisation and short term, risky debt. To do so, she collects data on two years, just before the Global Financial Crisis of 2008, on a number of firms from different sectors. She gets quite detailed information on such firms, namely the age, the board composition, how many people sit in the board, total investment, market capitalisation, amount of short term debt, total number of people employed, and even how many of the employees have higher degrees. Provided the wild differences across firms, she looks for a way to control for any characteristics, and thanks to her experience she looks at relative quantities rather than absolute ones.

Variables list:

- id: unique identifier
- sector: industrial sector
- year: year of the observation
- emp: number of employees
- age: years since incorporation
- profit: amount of net profits, millions
- inv: total amount of investment, millions
- mktcap: total market capitalisation, millions
- shortdebt: amount of short term debt, millions
- nboard: number of members in the board
- wboard: number of women in the board
- skemp: number of employees with higher degree

Case 4

A bank wants to understand what are the main drivers of default rates on loans. It has collected a large number of observations on a sample of its clients and hires an economist to carry out an empirical investigation on the issue. Besides basic demographic information (age, gender, education), it also has partial knowledge on the financial status of every one of its clients, including the amount they defaulted on, if ever. In detail, it knows approximately income, total wealth, and wealth that is liquid, i.e., assets that can be sold quickly. Thanks to transaction data from credit cards, it can also measure how much every individual spends in healthcare.

The bank hands the full dataset to the economist and she starts reviewing some literature on the topic, to find an appropriate method and to get a refresher on the main economic theories and empirical findings. It turns out that she lacks most relevant data to fit a full theoretical model and obtain exact estimates, but she can carry out a neat empirical work. In particular, she knows that when an individual finds himself in trouble, he decides to sell some assets to keep on paying his loan. On the other hand, factors like income and

total wealth are clear criteria that the bank uses to grant a loan, which is the necessary premise to have a loan to default on or not. The economist also considers that clients prioritise expenses, which is obvious for food and sheltering, for example.

In light of all these considerations she embarks in the empirical analysis, bearing in mind that defaulting on a loan depends on a number of factors filled with endogeneity and reverse causality.

Variables list:

- id: unique identifier
- age: age at the time of observation
- female: dummy, female = 0
- taxincome: amount of taxable income, thousands
- totwealth: amount of total wealth, thousands
- liqwealth: amount of wealth with high liquidity, thousands
- healthexp: amount of health expenditure
- edu: years of education
- amountdef: total amount of past defaults if any, thousands