Causal Data Analysis

Guidelines for the short empirical project

Teamwork. Economic analysis is typically carried out in teams. To get used to this, you will work in teams of 2 students.

Tasks

- Research outline (deadline: December 1, midnight)
- Short analysis (Deadline: December 15, midnight)
- Upload **one** copy of the documents to the class Moodle

Research Outline. The aim of the research outline is to let professors give feedback on the plan of the analysis. The document should be short (max one page) and contain the following:

- research question (what are the treatment and outcome variables)
- hypothesis, mechanism (what is the expected result and why)
- data
- econometric method (what kind of endogeneity problem it solves), regression equations
- extension

Short analysis. The aim of the analysis is to empirically study a well-defined causal relationship. We place quality over quantity: do a short, but well-done analysis. The length of the paper should not be more than five pages (1.5 spaced, standard 2.54 cm margins), including everything (exhibits, references etc.).

Use one of the methods we studied (not only an OLS). Do not use more than two regressions. One should be your baseline, the other an extension. The extension may be another method, another specification or the same method run on a different sample (e.g., males-females, small-large firms, industrial-service firms, different time periods). Essentially, it is up to you what extension you do.

We will grade your work based on the following:

- work with the data (checking data, description of your data)
- clarity of the question
- potential endogeneity problems, your identification assumption (you may give an example when it fails)
- interpretation of your results (maybe policy suggestions)
- outlook of your paper (clarity and formatting of the text, exhibits, use of sections, references)

Causal Data Analysis

Data. You may use any data you want. To ease the search for a suitable dataset, here is a list of individual and firm-level data which are available free of charge. Note that getting the data and understanding it may take time.

Individual data
European Social Survey: https://www.europeansocialsurvey.org/
Current Population Survey (US data on individuals): https://www.census.gov/programs-surveys/cps/data/datasets.html
China Health and Retirement Longitudinal Study:
https://charls.pku.edu.cn/en/Data/Harmonized_CHARLS.htm#
Global Entrepreneurship Monitor: https://www.gemconsortium.org/

Firm-level data
Business Environment and Enterprise Survey: https://www.beeps-ebrd.com/data/
Competitiveness Research Network: https://www.comp-net.org/
Word Management Survey: https://worldmanagementsurvey.org/