Meeting follow-up - 2

29/06/2020

Summary of the points:

1. Labor outcomes to study
2. To-do list
3. Timing
4. Doubts and possible problems

Labor outcomes to study

Strategy:

- Cities are units represented by their NUTS equivalent.
- Test the effect of the application in multiple labor-related outcome variables, mediated by pollution:
    - **ONLY NUTS 2 cities:** Hours worked per person,
    - Employment/unemployment,
    - Proportion of working age population,
    - Total population

To-do list

GET DATA

- ◯ Review that I have data on outcomes, treatment and mediator
    - ◯ Important / basic :
        - ◯ TREATMENT
            - ◯ Announcement date of LEZ for each German city ❗
                - ◯ Asked for that info to the German ministry, they should answer soon.
            - ✅ LEZ stage of each LEZ in Germany
            - ✅ Treatment status of every other European city to discard treated

- ◯ OUTCOMES
  - ✔ Employed population / economic sector (Employment)
  - ◯ Hours worked / economic sector (NUTS 2)
    - ✔ For all countries but France and Poland, getting it soon
  - ◯ Unemployment (NUTS 2 - only from 2006)
  - ✔ Working age population (age distribution)
  - ✔ Total population
- ◯ MEDIATOR
  - ◯ yearly pollution means
    - ◯ **(extra)** Number of peaks and "high pollution episodes"
      - ◯ Pollution measurements of all stations, daily data.
      - ◯ Figure out how I will count the number of peaks
    - ◯ yearly pollution means that are "population weighted", to know the real effect on the population
      - ✔ Grid of population density for all of Europe
      - ✔ Grid of PM2.5 pollution (yearly) in Europe
      - ✔ Grid of NO2 pollution (yearly) in Europe
      - ◯ (to calculate)
- ◯ Controls / secondary
  - ✔ Environmental plans (Germany + others Europe)
  - ◯ To control for absenteeism / temporality **(NUTS 2 cities)**
    - ✔ Economically active population by sex and Sector (N 2)
      - ✔ Share of men or women in the work market (N 2)
    - ◯ Share of households with elders / children
      - ✔ Proxy: Dependency ratio based on aggregate ages
    - ◯ Investment/capita (NUTS 2)
      - ✔ For all countries but France and Norway, getting that data soon
    - ◯ Wage levels or income levels (NUTS 2)
      - ✔ For all countries but France and Poland, getting it soon
  - ✔ mean age
  - ✔ Mortality and birth rate
  - ✔ Proportion of elderly population
  - ◯ **(extra)** Size of the zone compared to the size of the city
    - ✔ perimeter of zones

- [x] perimeter of cities
- [ ] (not yet calculated)

- [ ] Find the name of the city that had an specially bad effect (small city in the border)
- [ ] Review that all sources and statistics make sense in size and in interpretation.

CLEAN DATA, SUMMARY STATISTICS

- [ ] Get all data in one same spatial definition
- [ ] Summarize data
- [ ] Create a list of possible controls for each treated city
  - [ ] Bound overall differences in levels of predictive variables (maximum and minimum)
    - [ ] Restrict big relative or absolute differences in controls?
  - [ ] Restrict cities outside 50-100km of a LEZ AND outside the same NUTS2 region (+ unite the Rhine-Ruhr area)
- [ ] Review it is ready for analysis (clean and meaningful)
  - [ ] Re-run code to get data from Eurostat

START ANALYSIS

- [ ] Do Synthetic Control for some cities and each outcome
  - [ ] First check which combination of controls reduce pre-treatment RMSE and choose accordingly
  - [ ] Do robustness checks:
    - [ ] Sizable effect: check pre and post (RMSE)
    - [ ] Check if restricting controls further changes results (SUTVA)
    - [ ] No anticipation: do various pre-intervention placebo tests (maybe try the announcement/implementation of the first LEZ)
  - [ ] Try if I can differentiate the effects of different stages of LEZ (ATE for each period?)
- [ ] Try to perform and include in the final project:
  - [ ] Try to do MASC
  - [ ] Try to do a "Inverse" synthetic control for a couple untreated cities (Hamburg being the principal of them)
  - [ ] Try to create a webpage that displays the methodology, data and results in an interactive way.
- [ ] WRITE the results.

Timing

JULY

- 29-04: Finish to get data and start data cleaning
- 6-10: Finish data cleaning and hopefully do first analysis before last meeting. Look if ratio signal/data is good enough for Synthetic Control
- 13-18: Do analysis of Synthetic control and robustness tests, attempt to do MASC and alternative methods
- 20-25: Write most of the text, check for mistakes and have something solid.

AGOUST

- 27-01: Iteratively correct and improve text, try to do inverse synthetic control.
- 3-6: Review, re-read, and present

Doubts and possible problems

- Worst case scenario:
    - Size of heterogeneous shocks is too big, there is too much noise and an analysis is not possible:
        - Diff-in-Diff (pooling), Diff-in-diff with matching (more similar to Synthetic C.)
        - Do an alternative project

- Should I put everything in relative terms (rates) instead of overall values? This increases the common support.
    - Does this decrease its credibility to you?

- Keep the possibility of focusing in 3-4 cities instead of doing all.
    - I think this can make the final document easier to write and understand. Even if I do them all, focusing in some examples should be good. (Choose examples of different types of cities.)

- Announcement dates and biased results. If there are no significant effects in a pre-intervention placebo, then can I say there are no significant anticipation effects?

- Do you see other possible flaws or confounding factors?

Other questions:

- I have had some serious trouble moving my things out of my flat in London, with considerable stress and unability to focus. I am considering to ask for an extension to the deadline.

- I think a great proportion of work has gone into planning a good research strategy, would be a good idea to include a "Research memory" as an Annex?

- I could report, in an Annex, the results for GDP and productivity even if I don't claim causality for those outputs. If they are consistent with the rest of the discourse it might be a good thing to report.

Thanks for reading and your thoughts!