Discussion of the paper "The *power of the pen*: Political influences on the legislative procedure in Europe"

Bas Machielsen (discussant)

2023-11-06

Introduction

- Recapitulation of research question and research objectives:
  - "What is the effect of meetings between policy-makers and (representatives of) interest groups during the legislative procedure?
  - Meeting $\rightarrow$ Amendment $\rightarrow \Delta$ Benefit/cost trade-off
  - Theoretical framework:
    - New law project brings forth some costs and benefits
    - Interested parties bargain about the allocation of these costs/benefits to various groups
  - The *next* amendment can be interpreted as the *status quo* in this bargaining process
- I want to talk about: the empirical strategy, endogeneity, and the measurement and validation of some key variables

Empirical strategy

- Suggested approach in the presentation:

$$\text{Benefits}_{s,r,a,l,t} = \alpha_0 + \gamma \text{Meeting}_{s,r,a,l,t} + X'_{s,r,a,l,t}\beta + \epsilon_{s,r,a,l,t}$$

- Where Meeting_s =1 if stakeholder s met with any legislator l between project law r at time t and project law r at time $t-1$
- (I am wondering if we need the index a here, as amendments are just snapshots of law l at time t)
- More importantly, changes vs. levels: I think for each amendment, you ideally want to know either:
  - The *status quo*: the level of benefits / costs incurred by each of the stakeholders
  - This might be difficult to measure / requires more far-fetched assumptions on the abilities of the LLM
  - Therefore: *change* in the *status quo*: relative to the "first" draft text, what happened?

Empirical strategy: Independent variable

- Why no intensive margin/count variable in addition to a dummy?
  - There might be convex returns to lobbying
- Considering the wide variety of interested parties, maybe best to group them together
  - Dummy $= 1$ if interest group $j \in J$ has met with *rapporteur*
  - Only real or legal persons rather than abstract entities
- The approach implies that, if a stakeholder is successful at bargaining during a meeting, this will be reflected in the *next* amendment
  - Rather than spread out over more than one amendment
  - This is the possibly the source of finding the attenuated, lower bound-influence of lobbying

Empirical strategy: Dependent variable

- Maybe it might be interesting to look at more raw measures before investigating *benefits*
- Concretely: conditional on meeting occurring with stakeholder from category j:
  - Do we observe more changes in the project law than if there is no meeting?
  - Does the time to a new amendment decrease?
- Benefits or (Δ benefits) might also not be a dichotomous variable
  - For example, you might distinguish between *adding* and *deleting* text, and finding out the nature of what is added or deleted

Endogeneity

- Selection bias probably exists in an OLS approach: stakeholders select themselves into the treatment
  - Probably you do not have the intuitive counterfactual of interest yet: we want to know what would have happened compared to when the meeting did not take place *ceteris paribus*, but this is governed by choice
  - Meaning their potential outcomes probably not comparable to outcomes in the control group
- I think a possible IV strategy might focus on how the meetings are planned:
  - Are there queues for meetings?
  - Who allocates these? Are there also rejected meeting requests?
- Placebo test with an "unrelated" independent variable, e.g. meetings with secretary or with parties that have nothing to do with the law

Empirical Strategy: Fixed Effects

- Reminder: suggested approach in the presentation:

$$\text{Benefits}_{s,r,a,l,t} = \alpha_0 + \gamma \text{Meeting}_{s,r,a,l,t} + X'_{s,r,a,l,t}\beta + \epsilon_{s,r,a,l,t}$$

- a_0 contains a lot: what kind of fixed effects do you need?
- What are the *least important* of all of these dimensions?
  - I think: calendar time t
  - So maybe: investigate temporal heterogeneity later by splitting up the sample
- You are left with s, r, a, l
  - I think you definitely need law fixed-effects r
  - Potentially legislator fixed effects l
  - Law-amendment fixed effects a impossible because that is where your identifying variation comes from
  - For each stakeholder (group) j a separate meeting dummy / count

Manual Data Suggestion

- Approach now (GPT) is a bit arcane
- Manually needed data can be used to train a language model more explicitly
  - BERT or any other standard architecture + a classifier head
- Validation approach much needed:
  - Compare answers to manually classified amendments
  - Possibly: use stock price information of stakeholders to validate 'value' of amendment
  - Could also be a dependent variable:
    $P_j = \alpha_0 + \alpha_1$Amendment x Stakeholder j has Meeting $+$ α_2Amendment $+ \epsilon_i$