Designing quasi-experimental studies: some examples from peacebuilding policies

Cyrus Samii

Department of Political Science, Columbia University,
and MacMillan Center, Yale University

October 21, 2010

Overview

- Effects of policy interventions that are applied at the micro level and that we are not able to manipulate experimentally.

- Review design principles for quasi-experimental studies: identification, allocation, sample size.

- Show two examples from my own work.

Identification

Conditional independence is the formal condition, but how do you make that convincing?

- Condition on all identified confounders, and then claim that anything else must be "random."

- Take advantage of some *verifiable source* of (conditional) quasi-random assignment.

Sample allocation

Given an identification strategy, some observations are more useful than others. You want to allocate your sampling resources accordingly.

Sometimes you can easily know in advance:

- ► RDD: you want observations near the cutoff.
- ► Random encouragement design (IV): maximize power for estimating the ITT, assuming you know compliance rate.

Sometimes you cannot know in advance:

- ► Information you need to exploit conditional independence will be measured in survey.
- ► You can do preliminary analysis of auxiliary data to determine ratio of controls for each treatment, for example.

Sample size determination

Only once you've pinned down the identification strategy and the consequent sample allocation does sample size determination enter the picture. You can do this analytically or you can use simulation.

Illustration 1: Peacekeeping and "security bubbles"

Identification

- "Fog of peacekeeping" plus admin data that peacekeepers had.
- Assumed this information was sufficient to identify communities that were equally likely to be exposed.
- Coarsened exact matching to construct matching strata.

Sample allocation

- No. of exposed/non-exposed communities fixed by id strategy.
- How to allocate a given sample size over these strata?
- Just using a fixed number because of administrative ease.
- Optimally we would explore power of different allocations.

Sample size determination

- Simulated a DGP and studied power. Minimally adequate sample size of 17 HHs per community in 50 communities.

Illustration 2: Reintegration assistance

Identification

- ► Fortuitous accident: bureaucratic failure.
- ► But, affected individuals concentrated in one region
- ► Concerned about "incidental imbalance."

Sample allocation:

- ► Confounders data were to be measured with survey itself.
- ► Optimal allocation strategy would look at available data to determine a good ratio of unexposed to exposed.
- ► Not knowing that, we sought a representative sample.
- ► Saved by the fact that only $1/3$ were exposed, so we had ~ 2 unexposed for every exposed (110 exposed, 261 unexposed).

Sample size determination:

- ► Multipurpose survey, so it was a "representative" sample.
- ► Optimal design would have used optimal sampling ratio.

Conclusions

These are lessons learned by doing. The design of a quasi-experimental study:

- Specifies the identification strategy, nailing down one's understanding of the verifiable source of exogenous variation.

- Draws out implications for the optimal sampling allocation.

- Performs power analyses based on this allocation.