Overview of Econometric Approaches of Policy Evaluation

Xiaodong Gong
Structure of the talk

• The definition of policy evaluation in this literature;
• The Identification problem of policy evaluation;
• Introduction of some widely used approaches under various regimes;
• Conclusions.
What do I mean by policy evaluation?

- Estimation of the \textit{causal effect} of exposure of a set of units to a policy (a program, an intervention, or a treatment: T) on some outcome Y. In another words, comparison of the two outcomes for the same unit when exposed and when not exposed to the policy.

\[ D_i = Y_i(1) - Y_i(0) \]

- E.g., job training; change of leader; stimulus packages, ...

- Causality vs. correlation
Corr(cpi, rpx) = 0.91;
or,
cpi = 31.66 + 2.48 rpx
[-14.39] [29.74]
Identification problem: Fundamental problem of causal inference

• The problem is that we can at most observe one of the two outcomes because the unit can only be exposed to only one level of the treatment:

\[ Y_i = Y_i(T_i) = \begin{cases} 
Y_i(0), & \text{if } T_i = 0 \\
Y_i(1), & \text{if } T_i = 1 
\end{cases} \]

• The key is to find the counterfactual via appropriate control groups.

• How tricky the problem is depends especially upon the relation between the outcomes and the assignment to the policy.
The ‘simplest’ case---random assignment (Randomised experiments)

• The exposure to the policy is not affected by the outcomes and its probability is either independent of or a known function of the other observed factors.
The approach used for the ‘simplest’ case

- The (average) effect can simply be identified using cross-tabulation:
  \[ D = \bar{Y}(1) - \bar{Y}(0) \]

- Or equivalently, using simple regressions:
  \[ Y_i = \alpha + DT_i + \varepsilon \]

- Examples: State health insurance program (lottery); Access to the long-term reliable birth-control methods in Zambia; impact of background of judges on case outcomes.
Examples of complications

• Units are heterogeneous and the exposure to the policy and the outcomes may be affected by
  – observed characteristics in a way unknown by the researchers;
  – unobserved characteristics;

• No overlap between the exposed and unexposed units by the policy design;

• The policy effects are different across units;

• The outcomes of un-treated units are affected by those of the treated;
A bit more complicated case: selection on observables

- The exposure to the policy is still not affected by the outcomes but its probability is not a known function of the other observed factors anymore, which implies,

- Unconfoundedness: Conditional upon observed covariates there are no unobserved factors that are associated both with the exposure to the policy and with the potential outcomes.

\[ T_i \perp (Y_i(0), Y_i(1)) | X_i \]

- Overlap: the support of the conditional distributions of \( X_i \) given \( T_i = 0 \) overlaps completely with that of \( X_i \) given \( T_i = 1 \).
Commonly used methods

- **Multi-variate regressions**: suppose that the potential outcomes can be expressed as functions linear in parameters, then
  \[ Y_i = \alpha + DT_i + \beta X_i + \gamma T_i \cdot (X_i - \bar{X}) + \varepsilon \]

- **Propensity score**: sometimes, when the relation between the probability of exposure and the covariates are not linear, instead of covariates, estimated propensity scores are used in the regressions.

- **Matching**: the missing potential outcomes are calculated using only the outcomes of a few nearest neighbours of the opposite treatment group.
Even more complicated case---Selection on unobservable

• The exposure to the policy is affected by the outcomes or both of them are affected by some unobserved factors.

• E.g., return to education (ability); voting behaviour (ideological); unemployment (motivation) ...

• There is no unified set of methods for this except for a few special cases.
Instrumental Variable approach

• If there exists some *Instrumental Variable* (such as eligibility) that affects the exposure but not the outcomes, then the effect could be identified.

• The intuition is that from the variation in the exposure introduced by the instrument, the policy effect could be identified.
Random discontinuity

• Often arising from administrative decisions, the exposure to the policy is determined by the value of a predictor (the forcing variable, say $X_i$) being on either side of a common threshold.

• This generates a discontinuity in the conditional probability of receiving the policy as a function of this particular predictor. Consequently, any discontinuity of the conditional distribution of the outcomes as a function of $X_i$ at the threshold is interpreted as evidence of a causal effect of the policy.
PHI take-up and Medicare Levy Surcharge (Gong and Gao, 2013)
Difference-in-differences

• With more and more richer data available to the researchers, especially panel or repeated cross-sectional data, this method has been used very often.

• The simplest case is one where outcomes are observed for units from two (similar) groups over two time periods, and only one of the two groups are exposed to the policy in the second period.

• Difference-in-differences: The average gain over time in the non-exposed group is subtracted from the gain over time in the exposed group.
Effects of welfare reforms on labour supply of lone mothers (Gong and Breunig, 2013)

- Welfare reforms in 2004 and 2006 aimed to encourage women with young children to work;
- Single mothers vs single women without children;
Diff-in-diff estimates: 2004 reform led to an increase of hours of work among a small number of lone mothers through job changes.
Synthetic control methods

• When there are more than one control groups and arguably none of them is similar enough to the treated group, sometimes, an artificial control group that is more similar can be constructed.
Figure 1. Trends in per-capita cigarette sales: California vs. the rest of the United States.

Abadie et al (2010)
Figure 2. Trends in per-capita cigarette sales: California vs. synthetic California.
Simulation methods for ex ante evaluations

• Based upon structural models by making strong assumptions on, e.g., behaviour of the units;
• Often used for ex ante evaluations.
Example: childcare costs and female labour supply (Breunig and Gong, 2012, 2013)

• Step 1: estimate the preferences of married women with young children to describe their behaviour of labour supply and childcare demand in relation to wage, childcare costs, and so on;

• Step 2: simulate their behaviour change under different policy settings.
Conclusions

• The main issue about policy evaluation that concerns economists is to identify the causal effect of the policy when only one of the two potential outcomes is observed;

• Various methods that are associated with different assumptions can be used; but,

• One needs to bear in minds those assumptions when interpreting the results.