PHILOSOPHY OF ECONOMICS & POLITICS

LECTURE 12: EVIDENCE-BASED SOCIAL POLICY

DATE 21 JANUARY 2019

LECTURER JULIAN REISS
Today’s agenda

- Recall Michael Young: segregation works — obviously!
- But how do we know?
- Today we are going look at a development in recent economics and social science that is receiving a lot of (philosophical...) press: evidence-based social policy
- We’ll understand how exactly this movement understands evidence, why it does so, and what’s wrong with it
- If we have time, I’ll briefly introduce the four main strategies to deal with the so-called ‘problem of external validity’
Evidence-based social policy

- EBSP is a strong and highly acclaimed recent movement especially in the areas of education and development economics
What, then, is EBSP?

Basic idea: social policy decisions should be based on the ‘best available evidence’

That much is obvious, trivial even. But what’s the ‘best available evidence’? The hierarchy of evidence

It can be seen as promoting a stance similar to one of John Stuart Mill’s: experiments are the gold standard for causal inference (method of difference)
List of ‘What Works Centres’ in UK:

- National Institute for Health and Care Excellence (NICE; Health and social care)
- Sutton Trust/Educational Endowment Foundation (Educational achievement)
- College of Policing/What Works Centre for Crime Reduction (Crime reduction)
- Early Intervention Foundation (Early intervention)
- What Works Centre for Local Economic Growth (hosted by LSE, Arup, Centre for Cities; Local economic growth)
- Centre for Ageing Better (Improved quality of life for older people)
- What Works Centre for Wellbeing (Wellbeing)
EBSP: Examples from education

- RCTs have been performed to test claims about:
  - the role of technology in schools (‘computer-aided instruction’)
  - class size
  - school vouchers (allocation of vouchers by lottery in Bogotá, Colombia; Metco-programme in Boston)
  - compulsory attendance laws (in the form of ‘natural experiments’)
  - achievement incentives (‘cash for grades’)
  - macro education policy (e.g., the effects of large-scale expansion of higher education in the West Bank and Gaza Strip)
In favour of RCTs

- Observational studies always face the problem of confounders

- The ‘standard solution’ is to condition on background factors (socio-economic status, income, age, gender...)

- But: (a) Every economist comes up with her own list of favourite background variables; and (b) results tend to be highly sensitive to the exact choice

- Randomisation is often thought to solve this problem

- A randomised trial (generally speaking) also allows the blinding of researchers, which helps with various biases
Against RCTs

- The questions that can be addressed with RCTs are very narrow, in a twofold sense:
  - They concern only the ‘efficacy’ of policies
  - They identify an ‘average causal effect’
- We’d like to know tons of things besides efficacy: cost-benefit considerations, side effects, implementation, moral, cultural and political considerations...
- Average causal effects are potentially true of no-one (no woman has 1.66 children); a positive overall result might mean that the policy is extremely beneficial to a small minority but harmful to the majority
Against RCTs

- One main advantage of RCTs in medicine is the possibility of blinding.
- But that’s hardly an option in social science (you can’t hide from a subject that you’ve sold him a bed net!)
- Randomisation might introduce novel kinds of artefacts: risk-averse subjects might not like the fact that they’re playing a lottery and leave the experiment in order to obtain the treatment through a different route for sure; so the test population might be unrepresentative with respect to risk aversion.
Against RCTs

- The main problem is, however, that of external validity: there is no guarantee that the results generalise to other settings, and the way in which RCTs are set up makes this problem particularly pressing here.

- In general, every study is subject to external validity issues (for an analogy, think of animal studies – we never know in advance whether something that works in animals will also work in humans).

- But RCTs give us knowledge ‘cheaply’, which comes at a cost because we don’t know what to do with it.
  - Since we don’t know why a result holds, we have no basis for generalising it.
  - Results may be relative to a specific intervention.
  - There may always be ‘general equilibrium effects’.
In defence of RCTs

- Proponents of RCTs counter that:
  - Everyone faces the problem of external validity (true)
  - We shouldn’t give up the benefit of executing our own interventions (false – interventions may create new biases)
  - We should build on a secure foundation and go on from there (false – even if the basis were secure, what does it help if it’s irrelevant to the purpose at hand)

- This last point ignores the identification issues RCTs have; and it doesn’t address these questions:
  - How many RCTs do we need to generalise?
  - How similar to the target situation does our RCT have to be?
In sum...

- Some of the criticisms of the evidence-based policy movement are made by the same authors who criticised the natural experiments/instrumental variables movement (e.g., Deaton, Heckman)

- These authors, consequently argue *for more theory* in evidence-based policy

- Perhaps we don’t need more theory but *better knowledge of the causal mechanisms* that are responsible for outcomes?
External validity: Strategies

- The literature distinguishes **four strategies** to deal with the problem:
  - **Investigating mechanisms** (Dan Steel and others): if C causes E in the lab, and we have reason to believe that the mechanism responsible continues to hold, we have reason to believe that C causes E in the relevant policy setting
  - **Causal tendencies** (Cartwright): if C causes E in the lab and C has a stable tendency or capacity to bring about E, then C will continue to contribute towards E even in the presence of disturbances
  - **Engineering** (Guala): if C causes E in the lab and we build the policy situation in such a way that it mimics the lab closely, we have reason to believe that C causes E there too
  - **Field experiments** (List): if C causes E in the lab and continues to do so in a variety of field settings, we have reason to believe that C causes E in the relevant policy setting