Mixing econometrics and epidemiology

S.Geneletti

Imperial College Department of Epidemiology and Public Health
BIAS II project, NCRM, ESRC

09/03/2009

Motivation

- BIAS II project has strong focus in causal inference +
- My personal experience with econometric treatment effects during PhD and current work in epidemiology +
- NCRM collaborative meeting fund +
- ADMIN project (NCRM) expertise in econometrics (Lorraine Dearden) +
- LEMMA 2 (NCRM) project keen interest in causal inference =
- *Today's seminar series and tomorrow's brain-storming session*

What the Seminar is about

- Causal inference - methods to identify and techniques to estimate causal effects of interventions, treatments
- Both econometrics and epidemiology have frameworks
- There has been some cross-pollination but there is scope for more

Areas of potential cross-pollination

- Frameworks (potential outcomes, causal DAGs, decision theoretic)
- Methods (instrumental variables, regression discontinuity designs, structural marginal models)
- Estimators (ETT, ATNT, LATE etc)

What has stopped the cross-pollination?

- Epidemiologists learn about causal inference from statisticians
- Econometricians learn about causal inference from econometricians (Chicago school)
- Statistics and econometrics have very different traditions
- Epidemiology is typically interested in average effects over populations
- Comes from the focus on the effects of drug treatments and clinical trials
- Econometrics is interested in how different subgroups of the population behave when treated not in an average effect

Epidemiology - move towards public health interventions?

- Focus in epidemiology so far on average treatment effects
- Increasing interest towards preventive epidemiology - i.e. public health interventions
- Closer to the situation of labour/education econometrics
- Interest shifts away from average effect towards effects of intervention on subgroups
- More scope to borrow from econometrics both in terms of methods and estimators

Examples of public health interventions

- Smoking ban
- Food labelling
- Ban on vending machines in schools
  1. How will it affect those who used vending machines??
  2. Will not affect children who did not?
- Similar to compulsory schooling to the age of 16
  1. Only changes it for those who wanted to drop out
  2. Does not change those who were planning on it anyway
- Econometric treatment effects ETT, ATNT might be useful in the epidemiologic example
- Not all interventions are "blanket", e.g. GP recommendations

Why is cross pollination a good idea?

- Econometrics and epidemiology are both used to evaluate and instigate policy making
- They also share a common problem:
- **Inference about causal effects from observational data**
- This gives rise to two main issues:
  1. What causal inference framework to use
  2. How to deal with the problem of unobserved confounding
- Touch on former via latter

Why is cross pollination a good idea?

- Econometrics and epidemiology are both used to evaluate and instigate policy making
- They also share a common problem:
- **Inference about causal effects from observational data**
- This gives rise to two main issues:
  1. What causal inference framework to use
  2. How to deal with the problem of unobserved confounding
- Touch on former via latter

The problem of the unobserved confounder

- Confounding leads to differential selection of individuals into a study sample
- In econometrics this typically manifests itself as self-selection (Heckman has studied this at length)
- In epidemiology there are a number of confounders that are often not observed
- Maybe sample recruitment is different but the biases are similar and similar methods can be used across the board

The problem of the unobserved confounder

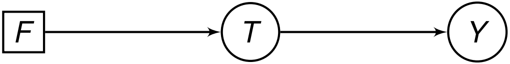
- Most methods developed in causal inference attempt to circumvent the problem of unobserved confounding
- Some methods imitate randomization
  - Instrumental variables
  - Regression discontinuity designs
- Other methods reweight data/parameters to recreate a “random” sample
  - Inverse probability weighting
  - Structural marginal models
- Focus on expressing the problem, talks will cover how to estimate!

The problem of the unobserved confounder

Notation for decision theoretic (Dawid 2001,2002 etc.)

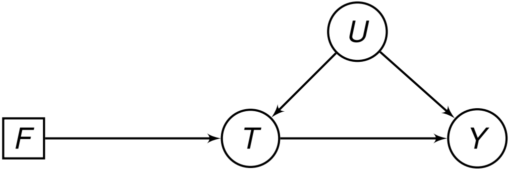
- Let Y be the outcome, and T the binary treatment and U the unobserved confounder
- Generally accept that unobserved confounders are not a problem when treatment is randomised
- They are a problem in observational studies
- We introduce a variable to represent this difference in "regimes" - F the intervention variable
- $p(T = t|F = t) = 1$ means set/forced $T = t$ as in a randomised trial
- $p(T|F = \emptyset) = p(T)$, T arises "naturally" in the observational study

Common problem continued



1. We can express the confounding problem in a DAG
2. Above means treatment assignment is *ignorable* ($Y \perp\!\!\!\perp F | T$)
3. However that does not tend to hold...

Common problem continued



1. We can express the confounding problem in a DAG
2. Above means treatment assignment is *ignorable* ($Y \perp\!\!\!\perp F | T$)
3. However that does not tend to hold...
4. Usually there is a confounder U ($U \perp\!\!\!\perp F$ and $Y \perp\!\!\!\perp F | (U, T)$)
5. If there is no randomisation estimating treatment effects becomes difficult

Defining the treatment effects

Decision theoretic

$$ATE_{dt} = E(Y|\mathbf{F}=\mathbf{1}, T=1) - E(Y|\mathbf{F}=\mathbf{0}, T=0)$$

not usually same as *naive treatment effect*

$$NTE_{dt} = E(Y|\mathbf{F}=\emptyset, T=1) - E(Y|\mathbf{F}=\emptyset, T=0)$$

unless treatment assignment is ignorable
Let

$$\begin{aligned} NTE_{dt}(u) &= E(Y|F=\emptyset, T=1, \mathbf{U}=\mathbf{u}) \\ &\quad - E(Y|F=\emptyset, T=0, \mathbf{U}=\mathbf{u}) \end{aligned}$$

Then

$$ATE_{dt} = \sum_U NTE_{dt}(u)$$

Defining the treatment effects

Potential outcomes notation

- T is binary treatment, Y is response
- $Y_t(i)$ is the response of i to treatment $T = t$
- If i receives $T = 1$, $Y_0(i)$ is counterfactual

Potential outcomes definitions

$$ATE_{pr} = E(Y_1 - Y_0)$$

- I'm not sure how to express NTE in this framework!?
- For more complex treatment effects different approaches diverge further

Final words

- Aims of today - to find areas of overlap and collaboration
- Showed some examples of epi-econ cross-pollination and scope for more
- Touched on the controversy in causal inference
- Used graphical models
- Mentioned methods that will be discussed later

Final words

Thanks for coming
I hope you enjoy the seminar!