CAUSAL EFFECT HETEROGENEITY IN OBSERVATIONAL DATA

ANIRBAN BASU
basua@uw.edu
@basucally

THE CHOICE INSTITUTE
School of Pharmacy

Background

> Generating evidence on effect heterogeneity in important to inform efficient decision making.
  - Translate to clinical decisions with sufficient reliability of evidence
  - Hypothesis generation for targeting future research
  - Creating algorithms for clinical decision support systems, and evaluation of CDSS
  - Making sub-group specific coverage decisions, where plausible
  - Appropriate value calculation (Today’s F4 session on curative therapies)
Background

> Reliable evidence -> accurate and unbiased

> Seek large samples for accuracy
> Seek some form of randomization for unbiasedness
> Seek cost-effective ways to produce such information

> Typical RCTs often fail on all aspects and are not the best mechanism to produce information on heterogeneity.
  – Usually do not have large sample sizes
  – Generalizability issues
  – Costly

OBSERVATION STUDIES TO RESCUE?
Exposure/Treatment

Observed Confounder
e.g. Age

Unobserved confounder:
Fundamental problem of evaluation
e.g. Disease Severity

Instrumental Variable

A good IV:
1) Strong predictor of Treatment choices (testable)
2) Is NOT independently associated with outcomes (untestable)
Exposure/Treatment

Unobserved confounder: Fundamental problem of evaluation

Instrumental Variable

Observed Confounder e.g. Age

A good IV:
1) Strong predictor of Treatment choices (testable)
2) Is NOT independently associated with outcomes (untestable)
What is an IV estimating?

- With a binary IV (e.g. two levels of formulary)
  - Local Average Treatment effect (Angrist and Rubin 1996)
  - Challenges:
    - Who are these people (remember we don’t observe some confounders in the data)?
    - How generalizable are there effects to other?
  - Partial salvation:
    - When the binary IV is a policy variable – LATE is at least interpretable
    - e.g. Oregon Medicaid Lottery
    - Better methods available with strong continuous instruments
**Advanced Econometric Methods**

**Employ an Economic Choice Model**

- Choice model tells us who is at the "margin" of choice
- Manipulation of IV levels help identify "marginal treatment effects" (MTEs)
- MTEs are building blocks for any interpretable mean treatment effect parameters
  - ATE
  - CATE
  - TT/TUT
  - PeT
Person-centered Treatment (PeT) Effects

> In a perfect RCT, one can estimate a
  – Population average treatment effect (pATE)
  – Conditional average treatment effect (CATE), e.g. the average effect of treatment for, say, 60-year old. → averages over all unobserved confounders

> With observational data, even with the same confounders measured, you can additionally learn about the unobserved confounder levels for a person by observing one’s choice and the circumstance (IV-level) in which the choice was made

> PeT effects are individualized effects conditioned on their observed confounder levels and averaged over their choice-specific unobserved confounder distribution.
  – Effect for each person in your sample, easily averaged over any factor

Empirical Example

Does transfer to intensive care units reduce mortality for deteriorating ward patients?
**Background**

- ICU Transfer versus stay in General Ward for hospitalized patients
- Prospective cohort study of the deteriorating ward patients referred for assessment for ICU transfer in the UK
- **Primary Outcome**: Death 7 days post assessment
- **Secondary Outcomes**: Death within 28 and 90 days
- **Baseline covariates**: demographics, some comorbidities, risk score
- **IV**: # of ICU beds available at the time of risk assessment. Vary across hospital and over time

**Average Effects**

<table>
<thead>
<tr>
<th></th>
<th>2 SLS</th>
<th>Bivariate probit</th>
<th>PeT Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-day mortality</td>
<td>-27.9%</td>
<td>-10.5%</td>
<td>-11.7%</td>
</tr>
<tr>
<td></td>
<td>(-73.8%, 18.0%)</td>
<td>(-47.1%, 26.2%)</td>
<td>(-25%, 1.5%)</td>
</tr>
<tr>
<td>28-day mortality</td>
<td>-34.0%</td>
<td>-7.9%</td>
<td>-4.9%</td>
</tr>
<tr>
<td></td>
<td>(-89.9%, 21.9%)</td>
<td>(-44.2%, 28.4%)</td>
<td>(-26.4%, 16.7%)</td>
</tr>
<tr>
<td>90-day mortality</td>
<td>-25.6%</td>
<td>-9.5%</td>
<td>-4.7%</td>
</tr>
<tr>
<td></td>
<td>(-83.8%, 32.5%)</td>
<td>(-48.1%, 29.1%)</td>
<td>(-28.5%, 19.2%)</td>
</tr>
</tbody>
</table>

- Notice the PeT estimates have narrower confidence intervals
Focus of 7-day mortality

Distribution of PeT Effects

Distribution of PeT effects

- All patients
- General ward
- ICU
- Critical care not recommended
- Critical care recommended
Individual Effect sizes

- PeT > -0.05 (N = 3952)
- -0.15 < PeT < -0.05 (N = 2553)
- PeT < -0.15 (N = 2508)

Size of circle represents number of patients
### Among those who transfer to ICU

#### Exploratory Multivariate Het. Prediction

|            | Coef.    | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|------------|----------|-----------|-------|-------|----------------------|
| logit_dead7 | Robust   |           |       |       |                      |
| age        | -.099409 | .0284813  | -3.49 | 0.000 | -.1552314            |
| age2       | -.0003175| .0002478  | -1.28 | 0.200 | -.0008033            |
| male       | -.4829999| .1071483  | -4.51 | 0.000 | -.6930068            |
| sepsis_dx  | .1307296 | .1519559  | 0.86  | 0.390 | -.1670985            |
| periarrrest| -.439148 | .4189958  | -1.05 | 0.295 | -.1.260365           |
| ccm1       | 1.74724  | .659889   | 2.62  | 0.009 | .419257              |
| ccm2       | -1.90807 | .6518933  | -2.93 | 0.003 | -.3.186594           |
| ccm3       | -4.687485| 1.344332  | -3.49 | 0.000 | -.7.32327            |
| news_score | -7.3532  | 1.009079  | -0.73 | 0.466 | -2.713079            |
| icnarc_score| -.7093541| .0508766  | -13.94| 0.000 | -.8090705            |
| sofa_score | -.1796798| .0155197  | -11.58| 0.000 | -.2100979            |
| out_of_hours| 1.26725  | .166926   | -16.08| 0.000 | -1.421747            |
| weekend    | 1.317777 | .1152948  | 11.43 | 0.000 | 1.091803             |
| winter     | .723735  | .2653278  | 2.73  | 0.006 | .2033406             |
| _cons      | 20.16815 | 1.383138  | 14.58 | 0.000 | 17.45725             |
Conclusions

> Application of novel econometrics methods to real-world data can be extremely productive
> Not all methods are created equal!
> Analysts need to weigh methods across domains of
  – causality,
  – interpretability,
  – precision,
  – ease of use
> Validation is a requirement for hypothesis generation exercises

References


Heckman JJ,Vytlacil EJ. Local instrumental variables and latent variable models for identifying and bounding treatment effects. Proc Nat Acad Sci 1999; 96(6): 4730-34


Improving Public Health Requires Inclusion of Underrepresented Populations in Research

Figure. Open NIH-funded Phase 3 and 4 studies as of October 19, 2017

- Pregnancy
- Lactation
- Child (<18 y)
- Older people (>65 y)
- Intellectual disability
- Physical disability

Jan-Feb; 77(1): 52–58

Figure 2.
Background

> Adult Intensive Care Units (ICU) costly and scarce resource
  - Supply usually lags demand
> No RCT evidence
> Observational study evidence
  - Do not deal with the endogeneity of transfer
  - Do not recognizing heterogeneity in returns from transfer
> Transfers to ICU typically relies on clinical judgement
  - Not perfect proxy for reliable and causal evidence

Our Study

> Exploit natural variation in ICU transfer according to ICU bed availability for deteriorating ward patients in the UK
> The (SPOT)light Study (N = 15,158)
  - Prospective cohort study of the deteriorating ward patients referred for assessment for ICU transfer
  - Hospitals were eligible for inclusion if they participated in the ICNARC Case Mix Programme
  - Patients recruited between Nov 1, 2010 - Dec 31, 2011 from 49 UK NHS hospitals
  - A variety of exclusion conditions were applied to identify deteriorating ward patients who are equipoised to be transferred to ICU
**Data**

- **Primary Outcome**: Death 7 days post assessment
- **Secondary Outcomes**: Death within 28 and 90 days
- **Exposure**: ICU transfer vs care on general wards
- **Baseline covariates**: Age, diagnosis of sepsis, peri-arrest, dependency at ward assessment and recommended level of care post assessment (4 levels) and three physiology measures
  - National Early Warning Score (NEWS): whether respiratory rate, oxygen saturations, temperature, systolic blood pressure, pulse rate, a level of consciousness vary from the norm,
  - the Sequential Organ Failure Assessment (SOFA), and
  - the ICNARC physiology score

**IV**

- **IV = NBA**: Vary across hospital and over time
- **Key Assumptions**:
  - NBA at ward patient’s assessment directly affects one’s probability of transfer to ICU
  - NBA unconditionally independent of mortality of patients