Instrumental variables to estimate heterogeneous treatment effects

Anirban Basu
*The Comparative Outcomes, Policy, and Economics (CHOICE) Institute*
*University of Washington, Seattle*

The Fifth Annual Global Health Economics Colloquium: The Economics of Vulnerable Populations, at Home and Abroad
Improving Public Health Requires Inclusion of Underrepresented Populations in Research

Figure. Open NIH-Funded Phase 3 and 4 Studies as of October 19, 2017

Percent of Trials

- Pregnancy
- Lactation
- Child (<18 y)
- Older people (>65 y)
- Intellectual disability
- Physical disability
Figure 2.
Fig 1.
A multilevel model of factors contributing to disparities in clinical trials.
Why do we care about representative enrollment?
Why do we care about representative enrollment?

HETEROGENEITY
Why do we care about representative enrollment?
## Solutions

### Pragmatic trials

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Assessment of Pragmatism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruitment of investigators and patients</td>
<td>To what extent are the participants in the trial similar to patients who would receive this intervention if it was part of usual care?</td>
</tr>
<tr>
<td>Eligibility</td>
<td>How much extra effort is made to recruit participants over and above what would be used in the usual care setting to engage with patients?</td>
</tr>
<tr>
<td>Recruitment</td>
<td>How different are the settings of the trial from the usual care setting?</td>
</tr>
</tbody>
</table>
Solutions

Pragmatic trials

If heterogeneity is a real concern in the exclusion of vulnerable population in trials, then why is obtaining the average effect from a population-representative study any better!

Table 1. Nine Dimensions for Assessing the Level of Pragmatism in a Trial, as Proposed in the Pragmatic–Explanatory Continuum Indicator Summary 2 (PRECIS-2) Tool.*

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Assessment of Pragmatism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recruitment of investigators and participants</td>
<td>To what extent are the participants in the trial similar to patients who would receive this intervention if it was part of usual care?</td>
</tr>
<tr>
<td>Eligibility</td>
<td>How much extra effort is made to recruit participants over and above what would be used in the usual care setting to engage with patients?</td>
</tr>
<tr>
<td>Setting</td>
<td>How different are the settings of the trial from the usual care setting?</td>
</tr>
</tbody>
</table>

*DILEMMA
Solutions

> Observational Studies
Solutions

> Observational Studies

Selection Bias
Unobserved confounder: Fundamental problem of evaluation

e.g. Disease Severity

Exposure/Treatment

Observed Confounder e.g. Age

BRAIN STROKE

Intracranial STROKE

Hemorrhagic STROKE
Unobserved confounder: Fundamental problem of evaluation

- e.g. Disease Severity

Instrumental Variable

Exposure/Treatment

- Strong predictor of Treatment choices (testable)
- Is NOT independently associated with outcomes (untestable)

A good IV:

1) Strong predictor of Treatment choices (testable)
2) Is NOT independently associated with outcomes (untestable)

Observed Confounder
- e.g. Age
Unobserved confounder:
Fundamental problem of evaluation

Inferior to

Instrumental Variable

Exposure/ Treatment

A good IV:
1) Strong predictor of Treatment choices (testable)
2) Is NOT independently associated with outcomes (untestable)
Unobserved confounder: Fundamental problem of evaluation

e.g. Disease Severity

Instrumental Variable

Exposure/Treatment

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

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)
  o Local Average Treatment effect (Angrist and Rubin 1996)

  o Challenges:
    • Who are these people (remember we don’t observe some confounders in the data)?
    • How generalizable are there effects to other?

  o Partial salvation:
    • When the binary IV is a policy variable – LATE is at least interpretable
    • e.g. Oregon Medicaid Lottery
    • But generalizability is still an issue
Advanced IV Methods

• Need Strong, Continuous/Multivalued IV variable

• Local instrumental variable approaches
  o Allows estimation of “Marginal Treatment effects”
    • individual who are at the “margin” of their choice
    • Indifferent between two options
    • Small perturbation in the IV move their choices

• Average individual’s observed choice-specific margins:
  o Allows computation of “Person-centered Treatment (PeT) effects”
  o PeT effect: an average treatment effect for each person in your sample

Heterogeneity in the impact of type of schooling on adult health and lifestyle

Anirban Basu\textsuperscript{a,b,\textdagger}, Andrew M. Jones\textsuperscript{c,d}, Pedro Rosa Dias\textsuperscript{e}

\textsuperscript{a} The Comparative Health Outcomes, Policy, and Economics (CHOICE) Institute, Department of Pharmacy, and Departments of Health Services and Economics, University of Washington, Seattle, USA
\textsuperscript{b} National Bureau of Economic Research, Cambridge, MA, USA
\textsuperscript{c} Department of Economics and Related Studies, University of York, Department of Economics, York, UK
\textsuperscript{d} University of Bergen and Centre for Health Economics, Monash University, Victoria, Australia
\textsuperscript{e} Imperial College Business School, London, UK
**Fig. 3.** Dependence between PaT effects of comprehensive schooling on depression at ages 23 (horizontal axis) and 42 (vertical axis).

**Dependence between effects on smoking at age 23 and depression at age 42**

Corr (95%CI): .22 (-.15, .59)

Corr (95%CI): .47 (.13, 0.81)
Does transfer to intensive care units reduce mortality for deteriorating ward patients?

O’Neil S, Grieve R, Basu A
Distribution of PeT effects
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.
Conclusions

> Advanced IV methods when applied to real-world data can address issues of generalizability and heterogeneity
  – Need strong continuous IVs.

> Can serve both hypothesis testing and hypothesis generation exercises for many subgroups including vulnerable populations.

> In health care good IVs can be found – even better if we can voluntarily create them within our systems.
Questions
backup
Empirical Example


Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Heterogeneity in the impact of type of schooling on adult health and lifestyle

Anirban Basu a,b,∗, Andrew M. Jones c,d, Pedro Rosa Dias e

a The Comparative Health Outcomes, Policy, and Economics (CHOICE) Institute, Department of Pharmacy, and Departments of Health Services and Economics, University of Washington, Seattle, USA
b National Bureau of Economic Research, Cambridge, MA, USA
c Department of Economics and Related Studies, University of York, Department of Economics, York, UK
d University of Bergen and Centre for Health Economics, Monash University, Victoria, Australia
e Imperial College Business School, London, UK
Fig. 3. Dependence between PeT effects of comprehensive schooling on depression at ages 23 (horizontal axis) and 42 (vertical axis).

Corr (95%CI): .22 (-.15, .59)

Dependence between effects on smoking at age 23 and depression at age 42

Corr (95%CI): .47 (.13, .81)
Empirical Example

HEALTH ECONOMICS
Published online 9 October 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/hec.2996

HETEROGENEITY IN ACTION: THE ROLE OF PASSIVE PERSONALIZATION IN COMPARATIVE EFFECTIVENESS RESEARCH

ANIRBAN BASU*\textsuperscript{a,*}, ANUPAM B. JENA\textsuperscript{b}, DANA P. GOLDMAN\textsuperscript{c}, TOMAS J. PHILIPSON\textsuperscript{d} and ROBERT DUBOIS\textsuperscript{e}

\textsuperscript{a}Department of Health Services, University of Washington, Seattle and the National Bureau of Economic Research, Cambridge MA, Seattle, WA, USA
\textsuperscript{b}Department of Health Care Policy, Harvard Medical School, Boston, MA, USA
\textsuperscript{c}Leonard D. Schaeffer Center for Health Policy and Economics, University of Southern California, California, USA
\textsuperscript{d}Irving B. Harris School of Public Policy Studies, University of Chicago, Chicago, USA
\textsuperscript{e}National Pharmaceutical Council, Washington D.C., USA
Table IV. Predicted hospitalizations in 12 months following initiation of atypical antipsychotic therapy under various therapeutic scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Average annual number of hospitalizations</th>
<th>% change from Status quo</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status quo</td>
<td>1.83 (1.81–1.85)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>All patients started on branded group of AADs</td>
<td>1.73 (1.59–1.87)</td>
<td>–5.5</td>
<td>0.15</td>
</tr>
<tr>
<td>All patients started on generic group of AADs</td>
<td>2.07 (1.91–2.23)</td>
<td>13.1</td>
<td>0.001</td>
</tr>
<tr>
<td>All patients started on optimal predicted therapy</td>
<td>1.32 (1.26–1.40)</td>
<td>–27.9</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AAD, atypical antipsychotic drug.
Notes: p-values reflect comparisons of average annual number of hospitalizations under various scenarios to status quo.
Branded group of AADs include ziprasidone, quetiapine, and aripiprazole. Generic groups of AADs include risperidone and olanzapine.
Empirical Example

Are Elderly Patients With Clinically Localized Prostate Cancer Overtreated? Exploring Heterogeneity in Survival Effects

Anirban Basu, PhD,* and John L. Gore, MD, MS†

Background: Clinical trial evidence shows minimal survival gains and higher complication rates from radical prostatectomy (RP) versus watchful waiting (WW) for elderly men with localized prostate cancer (PCa). It is believed that these patients are overtreated. The current analyses aim to explore patient-level heterogeneity in survival effects, examine matching of patients to treatments in practice, and identify patient characteristics driving heterogeneous effects, in order to present more comprehensive evidence about the concerns of overtreatment.

Methods: Eleven-year all-cause and PCa-specific survival among SEER-Medicare patients diagnosed during 1996–2002 were analyzed using local instrumental variable approaches.

Results: Of 201,930 men eligible for RP, 71.9% received RP in practice. Patients from RP were much more likely to receive RP in practice. Such positive self-selection was driven by PCa-specific survival than overall survival. Several comorbidities may play a critical role in predicting who could benefit from RP.

Conclusions: Our analyses corroborate concerns about PCa overtreatment. A small fraction of screen-detected PCa patients derive survival benefits from RP. Prediction tools should account for patient comorbidities to accurately predict survival benefits of RP over WW.

Key Words: prostate cancer, radical prostatectomy, survival, heterogeneity

(Med Care 2015;53: 79–86)
References


Heckman JJ, Vytlacil EJ. Local instrumental variables and latent variable models for identifying and bounding treatment effects. Proc Nat Acad Sci 1999; 96(8): 4730-34
