Identification of effect-heterogeneity using instrumental variables

Anirban Basu
University of Washington, Seattle and NBER
Symposium on OBSERVATIONAL STUDIES IN A LEARNING HEALTH SYSTEM
April 25, 2013
Acknowledgements

- Support for NIH grant RC4CA155809 and R01CA155329

- Support from AHRQ R13HS021019 that facilitated rich discussions on these issues.
Role of instrumental variables (IVs)

• Recap
  o IVs are variables that influence treatment choices but are independent of factors that determine potential outcomes.
  o IVs are viewed as natural randomizers
  o They can be used to establish causal treatment effects – taking into account both overt and hidden biases
(Stukel et al, 2007)

Treatment: Invasive cardiac treatment

Adjusted = 0.51
95% CI: 0.50-0.52

Prop. Score = 0.54
95% CI: 0.53-0.55

Prop.-based matching = 0.54
95% CI: 0.52-0.56

Outcome: Long-term AMI Mortality rate

Unobserved confounder:
Treatment selection of lower-risk patients

Observed confounders:
Age, sex, race, socio-economic status, comorbidities, inpatient treatments
Observed confounders:
- Age, sex, race, socio-economic status, comorbidities, inpatient treatments

Unobserved confounder:
- Treatment selection of lower-risk patients

Instrumental Variable
- Regional catheterization rate

Treatment
- Invasive cardiac treatment

Relative Rate = 0.84
95% CI: 0.79-0.90

Outcome
- Long-term AMI Mortality rate

(Stukel al, 2007)
Effect heterogeneity

• Effect heterogeneity complicates the evaluation problem. There is **no reason to believe** that:
  o the causal treatment effect estimated using an observational data has to match the average effect in an RCT.
  o the average treatment effect is a relevant metric for evaluation.
  o the IV effect has a relevant interpretation.
Outline

• Highlight how a choice model help in the interpretation of IV effects in the presence of heterogeneity.

• How local instrumental variable (LIV) methods help overcome problems with traditional IV approaches

• Person-centered Treatment (PeT) effects

• An empirical example
A choice model

- Consider a binary treatment, D.
- Assumption: Choice of D is based on an underlying latent index

\[ D = \begin{cases} 
1 & \text{if } U^* > 0 \\
0 & \text{if } U^* \leq 0 
\end{cases} \]

- A formal model for latent index

\[ U^* = h(X_o, Z) + \nu(X_u, \varepsilon) \]

- Observed confounders
- Instrumental variables
- Unobserved confounders
- Pure error
Data for $Z = -1, X_0 = -1$

<table>
<thead>
<tr>
<th>$Z$</th>
<th>$X_0$</th>
<th>$X_U$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-1</td>
<td>+1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$U^*$</th>
<th>$D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>-0.5</td>
<td>0</td>
</tr>
<tr>
<td>0.5</td>
<td>1</td>
</tr>
</tbody>
</table>
Interpretation of IV effect

• If treatment effects vary over unobserved confounders:
  o “essential heterogeneity” (Heckman & Vytlacil, 1999)
  o IV effect has generalizability issues
  o IV effect varies with specific IVs used
  o Alternative methods required to estimate population mean effect parameters: ATE, TT or TUT
  o Heckman & colleagues propose Local instrumental variable (LIV) methods
    • Help to identify and estimate Marginal Treatment Effects (MTEs)
Data for $Z = -0.5 - \varepsilon, X_0 = -1$

Data for $Z = -0.5, X_0 = -1$

Data for $Z = +1 - \varepsilon, X_0 = 1$

Data for $Z = +1, X_0 = 1$
MTEs

• Treatment effect for individuals at the margin of choice.
  o Conditional treatment effect for an individual with specific levels of $X_0$ and also $X_u$ defined by choice model given $X_0$ and $Z$.

• All population mean treatment effect parameters can be computed by aggregating MTE’s with varying weights.
Estimator

1\textsuperscript{st} stage

\[
\text{logit}(D) = \alpha_0 + \alpha_1 X + \alpha_2 Z
\]

2\textsuperscript{nd} stage

\[
E(Y \mid X, \hat{p}(x,z)) = g(\beta_0 + \beta_1 X + \beta_2 x \cdot \hat{p} + K(\hat{p}; \hat{\beta}_3))
\]

\[
MTE(x, U_D = u_D) = \left. \frac{d\hat{E}(Y \mid X, \hat{P}(x,z))}{d\hat{P}} \right|_{\hat{P}=p=(1-u_D)}
\]

\(U_D\) = Propensity for treatment selection based on unobserved confounders
Extension to Person-centered Treatment (PeT) effect

- PeT effect are conditional treatment effect that:
  - Not only conditions on observed risk factors
  - Averages over the conditional distribution of unobserved risk factors, conditioned on choices.

- Basu A. Journal of Applied Econometrics (Forthcoming)
Extension to Person-centered Treatment (PeT) effect

• PeT effect are conditional treatment effect that:
  o Not only conditions on observed risk factors
  o Averages over the conditional distribution of unobserved risk factors, conditioned on choices.

• They help to comprehend individual-level treatment effect heterogeneity better than CATEs.

• They are better indicators for the degree of self-selection than CATE.

• They can explain a larger fraction of the individual-level variability in treatment effects than the CATEs.

Basu A. Journal of Applied Econometrics (Forthcoming)
A Case Study for Evaluating Antipsychotic Drugs
Case of antipsychotics

- CATIE - found initializing with first vs second generation antipsychotics produced similar effectiveness over an 18-month period.
  - Lead to no major change in clinical practice (Chen et al., 2008)
  - Post CATIE, 40% of the state-run Medicaid programs have instituted prior authorization (PA) restrictions on some 2nd gen. drugs (Polinski et al, 2007)
  - Manufacturers are seen to have waning commitments to invest in developing new drugs in neuroscience (Reuters, Feb 2011)
Issues of Generalizability

• CATIE recruited patients, most were continuing to receive an AAD.
  o % with exacerbation of symptoms in past 3 mo = 28%
  o % not using any drug at baseline = 28%
• A first-line PA policy would apply to initiators or clean starters – do CATIE results apply?
• Compare with clean starters in Medicaid (no AAD in past 6 months)
  o % of patients hospitalized annually due to schizophrenia related symptoms: CATIE: 7-14%; MEDICAID: 36%
  o Avg. number of hospitalizations per patient in a year among those initiated with risperidone: CATIE: 0.20; MEDICAID: 0.75.
## RESULTS

**TABLE 2**: Predicted impact of generic group AADs compared to branded group AADs on average number of hospitalizations in 12 months following initiation of therapy

<table>
<thead>
<tr>
<th>Group</th>
<th>All hospitalizations Mean (95% CI)</th>
<th>Schizophrenia-related hospitalizations Mean (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients (ATE)</td>
<td>0.35 (0.02, 0.67)</td>
<td>-0.07 (-0.28, 0.10)</td>
</tr>
<tr>
<td>Patients initiating therapy with generic group (TT)</td>
<td>0.17 (-0.17, 0.44)</td>
<td>-0.15 (-0.38, -0.03)</td>
</tr>
<tr>
<td>Patients initiating therapy with branded group (TUT)</td>
<td>0.61 (0.29, 1.05)</td>
<td>0.002 (-0.13, 0.22)</td>
</tr>
<tr>
<td>TT - ATE</td>
<td>-0.18 (-0.13, -0.28)</td>
<td>-0.08 (-0.04, -0.12)</td>
</tr>
</tbody>
</table>

Note: IVs used are Provider’s (prescriber) rate of generic use and Zip-code specific rates of generic use. Both instruments are highly predictive of treatment choice, with a z > 9 for each in a logistic regression. F-statistics for Instruments > 600. They appear to achieve balance in observed confounders and also in the distribution of specific drugs within generic or branded groupings.
Treatment effect heterogeneity

PeT Effects of Generic Group vs Branded Group of AADs

On # of Schizophrenia-related hospitalizations in Year 1

<table>
<thead>
<tr>
<th>On # of Overall hospitalizations in Year 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
</tr>
<tr>
<td>-5</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>25</td>
</tr>
</tbody>
</table>

20% 21%
48% 11%
What explains treatment effect heterogeneity?

**TABLE 3:** Percent variation in PeT effects explained by subgroups

<table>
<thead>
<tr>
<th>Sub-group</th>
<th># of subgroups</th>
<th>Variance in treatment effects explained by subgroups (%) for All hospitalizations</th>
<th>Variance in treatment effects explained by subgroups (%) for Schizophrenia-related hospitalizations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>2</td>
<td>1.2</td>
<td>1.0</td>
</tr>
<tr>
<td>Race</td>
<td>4</td>
<td>4.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Age</td>
<td>47</td>
<td>4.2</td>
<td>2.1</td>
</tr>
<tr>
<td>Total number of comorbidities</td>
<td>15</td>
<td>46.7</td>
<td>18.4</td>
</tr>
<tr>
<td>Unique combinations of comorbidities</td>
<td>5,905</td>
<td>87.8</td>
<td>67.1</td>
</tr>
</tbody>
</table>

In other words…
A CER trial with 6,000 treatment groups needed explain 88% of variation in effects.
Selection in practice

PeT Effects of Generic Group vs Branded Group of AADs

On # of Schizophrenia-related hospitalizations in Year 1

%

%

%

%
## POLICY IMPLICATIONS

**TABLE 4:** 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 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All patients started on branded group of AADs</td>
<td>1.73 (1.59 – 1.87)</td>
</tr>
<tr>
<td>All patients started on generic group of AADs</td>
<td>2.07 (1.91 – 2.23)</td>
</tr>
</tbody>
</table>

Notes: P-values reflect comparisons of average annual number of hospitalizations under various scenarios to status quo.
**POLICY IMPLICATIONS**

**TABLE 4:** 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 (95% CI)</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>
</tbody>
</table>

Notes: P-values reflect comparisons of average annual number of hospitalizations under various scenarios to status quo.
**POLICY IMPLICATIONS**

**TABLE 4:** 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 (95% CI)</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>

Notes: P-values reflect comparisons of average annual number of hospitalizations under various scenarios to status quo.
Implications for CER

• Individual patient differences of crucial importance
  o Tying decision making – i.e., coverage or clinical guidelines –
    to average results may be marginally beneficial at best,
    substantially harmful at worst
  o Studying heterogeneity over broad subgroups may not be
    useful

• Learning-by-doing works well, in some contexts

• Algorithmic predictions may be a promising way to guide clinical decision-making
  o Large, observational data sets can be valuable resources to
    explore and generate such algorithms, which can be later
    validated using confirmatory studies
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
