Statistical Methods III: Spring 2013

Jonathan Wand

Stanford University

Choice + Inference

Outline

1. Choice model with discrete endogenous variables: LATE

Get-out-the-Vote Messages (GOTV)

Question: Is it possible to increase the likelihood of an individuals turnout by making an appeal to vote?

Features of observational studies

- Contact may be correlated with outcome
    - Candidates may be more likely to target individuals who they think will turnout
    - Politically active individuals may be more likely to be in contact with a candidate and also more likely to value voting
- Difficult to measure the quality/quantity of contact and mobilization.
    - Often rely on self-reports/memory recall of citizens
    - More effective contacts will be remembered

Get-out-the-Vote Messages (GOTV)

Question: Is it possible to increase the likelihood of an individuals turnout by making an appeal to vote?

Features of experiment

- control who is assigned to "treatment" or not
- control quality/quantity of "treatments"

Gerber and Green (1998) field experiment sample of households with one or two individuals registered to vote:

- random assignment to zero, one, or more types of treatments:
  - in-person contact
  - telephone call
  - direct mail
- random assignment to appeal
  - civic duty
  - close election
  - neighborhood solidarity

GOTV design

Design for in-person contact:

Group	Number of people
Treatment	5,800
Control	23,500

Variable	Description
$D_i \in \{0, 1\}$	i treated / contacted
$Y_i \in \{0, 1\}$	Outcome : Voted or not
$Y_i(1)$, $Y_i(0)$	Potential outcome under contact or not

What we may want to learn, ATE $E[Y_i(1) - Y_i(0)]$.

- If $D_i \perp\!\!\!\perp Y_i(0), Y_i(1)$ then $E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)]$

Get-out-the-Vote (GOTV)

From Gerber and Green (GG, 1998), “in-person” RCT:

Group	N	Type
Treatment assigned and received	1,600	compliers
Treatment assigned **not** received	4,200	non-compliers
Control group	23,500	compliers

In-person Contact rate: **28 percent**

With non-compliance, we have a problem similar to observational study:

- those who receive treatment may not be random
- particular fear:
  individuals who are are more likely to be contacted may also be more likely to vote

Angrist, Imbens, Rubins

AIR notation	Description
$D_i(Z) \in \{0, 1\}$	Contact given assignment Z
$Y_i(Z, D) \in \{0, 1\}$	Voted given assignment and actuality

Potential Treatment	Description
$D_i(Z_i = 1)$	Assigned to treatment
$D_i(Z_i = 0)$	Assigned to control

Potential Outcome	Description
$Y_i(Z_i = 1, D_i = 1)$	Assigned to treatment, and treated
$Y_i(Z_i = 1, D_i = 0)$	Assigned to treatment, and *not* treated
$Y_i(Z_i = 0, D_i = 1)$	Assigned to control, and treated
$Y_i(Z_i = 0, D_i = 0)$	Assigned to control, and *not* treated

Only going to see one outcome for an individual.

Difference in observed outcomes - randomized

What does simple randomization give us?

$$Y_i(0), Y_i(1) \perp\!\!\!\perp D_i$$

therefore,

$$E[Y_i(0) \mid D_i = 0] = E[Y_i(0) \mid D_i = 1] = E[Y_i(0)]$$
$$E[Y_i(1) \mid D_i = 0] = E[Y_i(1) \mid D_i = 1] = E[Y_i(1)]$$

Thus, no selection bias,

$$0 = E[Y_i(0) \mid D_i = 1] - E[Y_i(0) \mid D_i = 0]$$

and thus,

$$\begin{aligned} & E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0] \\ &= E[Y_i(1) - Y_i(0) \mid D_i = 1] + \{E[Y_i(0) \mid D_i = 1] - E[Y_i(0) \mid D_i = 0]\} \\ &= E[Y_i(1) - Y_i(0) \mid D_i = 1] \\ &= E[Y_i(1) - Y_i(0)] \end{aligned}$$

Non-compliance as an identification problem

Without compliance ($D_i = Z_i$), potential confoundness,

$$Y_i(0), Y_i(1) \not\perp\!\!\!\perp D_i$$

if so, then

$$E[Y_i(0) \mid D_i = 0] \neq E[Y_i(0) \mid D_i = 1] \neq E[Y_i(0)]$$
$$E[Y_i(1) \mid D_i = 0] \neq E[Y_i(1) \mid D_i = 1] \neq E[Y_i(1)]$$

Thus, potential selection bias,

$$0 \neq E[Y_i(0) \mid D_i = 1] - E[Y_i(0) \mid D_i = 0]$$

and thus,

$$E[Y_i \mid D_i = 1] - E[Y_i \mid D_i = 0]$$
$$\neq E[Y_i(1) - Y_i(0) \mid D_i = 1] + \{E[Y_i(0) \mid D_i = 1] - E[Y_i(0) \mid D_i = 0]\}$$
$$\neq E[Y_i(1) - Y_i(0) \mid D_i = 1]$$
$$\neq E[Y_i(1) - Y_i(0)]$$

Non-compliance, overcoming identification problems

Approaches to identification with non-compliance

1. Bounds (cf, Manski)
2. Parametric/structural (cf, Heckman)
3. Redefine estimand of interest
    1. Intention-to-treat (ITT) for outcome
       With

       $$Z_i \perp\!\!\!\perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$$

       we can estimate

       $$\text{ITT} = E[Y_i(1, D_i(1)) - Y_i(0, D_i(0))] = E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]$$

    2. Could also estimate ITT for receiving treatment

       $$\text{ITT}_t = E[D_i(Z_i = 1) - D_i(Z_i = 0)] = E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0]$$

    3. Local Average Treatment Effect (LATE)

What do we need for LATE?

- IA refer to assignment indicator Z as an "instrument"
- interpretation hinges on properties holding..
    1. instrument exists
    2. monotonicity holds
- Two contrast points from parametric model to think about
    1. what parameter are we identifying?
    2. what assumptions do we need?

LATE

Condition 1 (Existence of Instruments)

(a) Joint independence

$$Z_i \perp\!\!\!\perp (Y_i(0), Y_i(1), D_i(0), D_i(1))$$

- adds another restriction to $D_I \perp\!\!\!\perp Y_i(0), Y_i(1)$
- testable?

(a') Exclusion, for all z, z', d',

$$Y_i(d) = Y_i(z, d) = Y_i(z', d)$$

- interpretation?
- cf. AIR has good discussion

(b) Non-trivial effect of assignment

$$E[D_i \mid Z_i = 1] \neq E[D_i \mid Z_i = 0]$$

- testable?

LATE

Condition 2 (Monotonicity)

$$D_i(1) \geq D_i(0)$$

If we had this equation:

$$D_i(z) = 1\{\gamma_0 + z_i\gamma_1 + \epsilon_i > 0\}$$

what would we need to assume for monotonicity to hold?

Defining (non)compliance

	$D_i(0) = 0$	$D_i(0) = 1$
$D_i(1) = 0$	never-taker	defier
$D_i(1) = 1$	complier	always-taker

- if all compliers, then

$$D_i(Z_i) = Z_i$$

- for all except for defiers, we have,

$$D_i(Z) \leq Z$$

- for Defiers,

$$D_i(Z) > Z$$

- Q: do we get observe which cell we are in? i.e., type?

LATE

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	complier/never-taker	never-taker/defier
$D_i = 1$	always-taker/defier	complier/always-taker

- If $Z_i = 0$ and $D_i = 0$, then could be either complier OR never-taker
- If $Z_i = 0$ and $D_i = 1$, then could be either never-taker OR defier
- and so on...
- LATE needs absense of defiers

Conditional probabilities

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	$\pi_c + \pi_0$	$\pi_0 + \pi_d$
$D_i = 1$	$\pi_1 + \pi_d$	$\pi_c + \pi_1$

- Each cell is $E(D_i = w \mid Z = z)$
- If no defiers, then proportion of each type is identifiable from this table of conditional probabilities

Conditional probabilities

	$Z_i = 0$	$Z_i = 1$
$D_i = 0$	$\pi_c + \pi_0$	$\pi_0 + 0$
$D_i = 1$	$\pi_1 + 0$	$\pi_c + \pi_1$

- Each cell is $E(D_i = w \mid Z = z)$
- If no defiers, then proportion of each type is identifiable from this table of conditional probabilities,

Population parameters,

$$\pi_1 = E(D_i \mid Z_i = 0) = P(always)$$
$$\pi_0 = 1 - E(D_i \mid Z_i = 1) = P(never)$$
$$\pi_c = E(D_i \mid Z_i = 1) - E(D_i \mid Z_i = 0) = P(complier)$$
$$\pi_d = 0 = P(defier)$$

Refresher: Law of Total Probability

Can partition any single event into multiple disjoint events

$$E = (E \cap F) \cup (E \cap F^c).$$

I.e., E can occur in two mutually exclusive ways:

$$\begin{aligned} P(E) &= P\big((E \cap F) \cup (E \cap F^c)\big) \\ &= P(\ E \cap F) + P(E \cap F^c) && \text{(why?)} \\ &= P(E \mid F)P(F) + P(E \mid F^c)P(F^c) && \text{(why?)} \end{aligned}$$

Definition (Law of Total Probability)

Given events $E, F \in \Omega$,

$$P(E) = P(E \mid F)P(F) + P(E \mid F^c)P(F^c).$$

Refresher: Iterated Expectations

Theorem (Law of iterated expectations)

If X and Y are any two random variables then

$$E_X X = E_Y \left[\, E_{X|Y}(X \mid Y) \, \right]$$

Proof:

$$E_X X = \sum_x \sum_y x \, f(x, y)$$

Imbens-Angrist Decompositions

Conditional expectation by types of respondents:

$$
\begin{aligned}
E[Y_i \mid Z_i = 1] \quad = \quad & E[Y_i \mid Z_i = 1, \text{complier}] \times P(\text{complier}) + \\
& E[Y_i \mid Z_i = 1, \text{never}\ ] \times P(\text{never}\ ) + \\
& E[Y_i \mid Z_i = 1, \text{always}\ ] \times P(\text{always}\ ) + \\
& E[Y_i \mid Z_i = 1, \text{defier}\ ] \times P(\text{defier}) \\
= \quad & E[Y_i(1) \mid \text{complier}] \times P(\text{complier}) + \\
& E[Y_i(0) \mid \text{never}\ ] \times P(\text{never}\ ) + \\
& E[Y_i(1) \mid \text{always}\ ] \times P(\text{always}\ ) + \\
& E[Y_i(0) \mid \text{defier}\ ] \times P(\text{defier}) \\
= \quad & E[Y_i(1) \mid \text{complier}] \times \pi_c + \\
& E[Y_i(0) \mid \text{never}\ ] \times \pi_0 + \\
& E[Y_i(1) \mid \text{always}\ ] \times \pi_1
\end{aligned}
$$

Imbens-Angrist Decompositions

Conditional expectation by types of respondents:

$$
\begin{aligned}
E[Y_i \mid Z_i = 0] \quad = \quad & E[Y_i \mid Z_i = 0, \text{complier}] \times P(\text{complier}) + \\
& E[Y_i \mid Z_i = 0, \text{never}] \times P(\text{never}) + \\
& E[Y_i \mid Z_i = 0, \text{always}] \times P(\text{always}) + \\
& E[Y_i \mid Z_i = 0, \text{defier}] \times P(\text{defier}) \\
\\
= \quad & E[Y_i(0) \mid \text{complier}] \times P(\text{complier}) + \\
& E[Y_i(0) \mid \text{never}] \times P(\text{never}) + \\
& E[Y_i(1) \mid \text{always}] \times P(\text{always}) + \\
& E[Y_i(1) \mid \text{defier}] \times P(\text{defier}) \\
\\
= \quad & E[Y_i(0) \mid \text{complier}] \times \pi_c + \\
& E[Y_i(0) \mid \text{never}] \times \pi_0 + \\
& E[Y_i(1) \mid \text{always}] \times \pi_1
\end{aligned}
$$

LATE

So, given these representations of population values,

$$E[Y_i \mid Z_i = 1]$$
$$= E[Y_i(1) \mid \text{complier}]\pi_c + E[Y_i(0) \mid \text{never}\,]\pi_0 + E[Y_i(1) \mid \text{always}\,]\pi_1$$

$$E[Y_i \mid Z_i = 0]$$
$$= E[Y_i(0) \mid \text{complier}]\pi_c + E[Y_i(0) \mid \text{never}\,]\pi_0 + E[Y_i(1) \mid \text{always}\,]\pi_1$$

We can solve for the difference

$$E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]$$
$$= E[Y_i(1) \mid \text{complier}]\pi_c - E[Y_i(0) \mid \text{complier}]\pi_c$$
$$= E[Y_i(1) - Y_i(0) \mid \text{complier}]\pi_c$$

LATE

Given difference

$$\begin{aligned} & E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0] \\ & \quad = E[Y_i(1) \mid \text{complier}]\pi_c - E[Y_i(0) \mid \text{complier}]\pi_c \\ & \quad = E[Y_i(1) - Y_i(0) \mid \text{complier}]\pi_c \end{aligned}$$

We can solve for,

$$E[Y_i(1) - Y_i(0) \mid \text{complier}] = \frac{E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]}{\pi_c}$$

Can we estimate π_c?

Yes, it is ITT for receiving treatment.

$$E[Y_i(1) - Y_i(0) \mid \text{complier}] = \frac{E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]}{E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0]}$$

LATE applied to GOTV

Vote Percent for Treatment: 47.2%; for Control: 44.8%
Percent Contacted of Assigned Treatment: 27.9%

Local Average Treatment Effect

$$E\Big[Y_i(1) - Y_i(0) \mid D_i(1) - D_i(0) = 1\Big] = \frac{E\Big[Y_i\big(1, D_i(1)\big) - Y_i\big(0, D_i(0)\big)\Big]}{E\Big[D_i(1) - D_i(0)\Big]}$$

$$= \frac{\text{ITT Vote}}{\text{ITT Contact}}$$

$$= \frac{.472 - .448}{.279} = .087$$

LATE

Questions to think about,

- instruments easy with randomized assignment; do you believe them in obs research?
- without control of experiment, is monotonicity plausible?
- how do assumptions differ from parametric model?
- what is weird about LATE?
- how does value depend on instrument?