INSTITUTE OF DEVELOPING ECONOMIES

Discussion Papers are preliminary materials circulated to stimulate discussions and critical comments

DISCUSSION PAPER No.106

Bounding ATE with ITT

Seiro Ito[†]

May 2007

Abstract We propose the bounds on ATE using intention-to-treat (ITT) estimator when there are compliers/noncompliers in randomized trials. The bounds are given as ITT⩽ATE⩽CACE, where compliers' average treatment effect (CACE) can be computed from ITT and complier ratio. We show that these bounds can be derived from two assumptions: (1) average treatment effect is greater with compliers than noncompliers or CACE⩾NACE, (2) noncompliers' average treatment effect (NACE) is nonnegative. We give an example of poverty impacts of health insurance, and effects of adverse selection and moral hazard of health insurance.

Keywords intention-to-treat (ITT) estimator, compliers' average treatment effect (CACE) estimator, ATE estimator, bounds

JEL classification C13, C93, D82, I11, O15

[†] Institute of Developing Economies

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

Bounding ATE with ITT

[Very Preliminary: Comments welcome.]

May 21, 2007

Seiro Ito[†]

Abstract We propose the bounds on ATE using intention-to-treat (ITT) estimator when there are compliers/noncompliers in randomized trials. The bounds are given as ITT⩽ATE⩽CACE, where compliers' average treatment effect (CACE) can be computed from ITT and complier ratio. We show that these bounds can be derived from two assumptions: (1) average treatment effect is greater with compliers than noncompliers or CACE⩾NACE, (2) noncompliers' average treatment effect (NACE) is nonnegative. We give an example of poverty impacts of health insurance, and effects of adverse selection and moral hazard of health insurance.

Keywords intention-to-treat (ITT) estimator, compliers' average treatment effect (CACE) estimator, ATE estimator, bounds

JEL classification C13, C93, D82, I11, O15

I Randomization and Intention-to-Treat Estimator

Consistent estimates of the treatment effect parameter of interest can be obtained by randomization of treatment. Treatment randomization, however, is rarely implemented in practice because it implies that the treatment must be provided to the participant, regardless of what he/she wishes. Researchers thus generally resort to the randomization of eligibility to treatment.

What can be estimated with eligibility randomization is the intention-to-treat (ITT) estimator. It is the difference in the outcome y between the eligible and the ineligible. ITT, however, is in general not equal to the average treatment effect (ATE), the most popular parameter in the treatment effect literature. Thus it is useful if we understand the relationship between ITT and ATE in an operational way. The purpose of this paper is to present such a relationship in terms of bounds of ATE.

If we are to estimate the mean impact, ITT estimator is given by:

$$ITT = \mathcal{E}[y|z = 1] - \mathcal{E}[y|z = 0].$$

We denoted $\mathcal{E}[\cdot]$ as an expectation operator, the eligible with $z = 1$ and the ineligible with $z = 0$. No covariates are included in the above, but they can be introduced in the analysis without complications.

Following Imbens and Rubin (1997), we classify the subject as: compliers w_c, defiers w_d, never-takers w_n, and always-takers w_a. We assume for simplicity the homogeneity of individuals among each group. Then we have:

$$\mathcal{E}[y|z = 1] = w_c y_{c1} + w_d y_{d0} + w_n y_{n0} + w_a y_{a1},$$
$$\mathcal{E}[y|z = 0] = w_c y_{c0} + w_d y_{d1} + w_n y_{n0} + w_a y_{a1},$$

where y_{i1} is group i's y with treatment, and y_{i0} is the outcome y of group i without treatment.

[†] Institute of Developing Economies

ITT is:

$$ITT = w_c(y_{c1} - y_{c0}) + w_d(y_{d0} - y_{d1}). \tag{1}$$

The most popular estimator of treatment effect, average treatment effect (ATE), and average treatment effect on the treated (ATE_1) are given by:

$$\begin{aligned} ATE &= w_c(y_{c1} - y_{c0}) + w_d(y_{d1} - y_{d0}) + w_n(y_{n1} - y_{n0}) + w_a(y_{a1} - y_{a0}), \\ ATE_1 &= \omega_c(y_{c1} - y_{c0}) + \omega_d(y_{d1} - y_{d0}) + \omega_a(y_{a1} - y_{a0}), \end{aligned}$$

where the weights ω_i, $i = c, d, a$ are conditional probabilities and are given as:

$$\omega_c = \tfrac{w_c}{q}, \quad \omega_d = \tfrac{w_d}{q}, \quad \omega_a = \tfrac{w_a}{q}, \quad q = pw_c + (1-p)w_d + w_a.$$

We can see that ITT is not equal to ATE nor ATE_1 in general. ITT will be equal to ATE only when $w_c = 1$, $w_i = 0$, $i = d, n, a$.

II Bounding ATE with ITT

We will set the bounds of ATE using ITT. For simplicity, we assume $w_d = w_a = 0$, $w_n = 1 - w_c$. Denote compliers' average causal effect (CACE) and noncompliers' average causal effect (NACE)[*1] as below:

$$\begin{aligned} CACE &= \mathcal{E}[y_1 - y_0 | i = c], \\ NACE &= \mathcal{E}[y_1 - y_0 | i = n]. \end{aligned}$$

Then ITT is:

$$\begin{aligned} ITT &= \mathcal{E}[y|z = 1] - \mathcal{E}[y|z = 0], \\ &= w_c\mathcal{E}[y|z = 1, i = c] + w_n\mathcal{E}[y|z = 1, i = n] - w_c\mathcal{E}[y|z = 0, i = c] - w_n\mathcal{E}[y|z = 0, i = n], \\ &= w_c\mathcal{E}[y_1 - y_0|i = c], \\ &= w_c CACE. \end{aligned}$$

ATE is:

$$ATE = w_c\mathcal{E}[y_1 - y_0|i = c] + w_n\mathcal{E}[y_1 - y_0|i = n] = ITT + w_n NACE.$$

The only unobservable term in the above is $\mathcal{E}[y_1|i = n]$, the mean counterfactual outcome of the noncompliers. We will assume that:

$$NACE \leqslant CACE. \tag{A1}$$

Possible justification behind this assumption is that individuals face the uniform opportunity cost in participation, and the self-selection process can be explained only with the individual gross benefit or the individual treatment effect $y_1 - y_0$. Since ATE is a weighted average of NACE and CACE, it is immediate that:

$$NACE \leqslant ATE \leqslant CACE.$$

The problem with these bounds is that we cannot observe NACE. So one needs to set the lower bound on NACE:

$$NACE \geqslant 0. \tag{A2}$$

[*1] They could have been termed as compliers'/noncompliers' average treatment effect, but we will follow the convention in the literature.

Possible justification can be that a lab experiment or other prior belief shows that treatment can not do any harm in gross outcome. These (A1) and (A2) are fairly weak assumptions. (A2) implies:

$$ATE = ITT + w_n NACE \geqslant ITT.$$

So we have:

$$ITT \leqslant ATE \leqslant CACE, \tag{2}$$

or

$$w_c \mathcal{E}[y_1 - y_0 | i = c] \leqslant ATE \leqslant \mathcal{E}[y_1 - y_0 | i = c]. \tag{3}$$

Width of the bounds is $(1 - w_c)CACE$, thus larger the w_c, smaller the width.

III Examples: Microinsurance

III.1 Setup

Suppose that an NGO in India is selling insurance policies to the poor. We would like estimate the impacts of adverse selection and moral hazard, separately from each other, on health care utilization y (number of doctor visits). Needless to say, it is important to distinguish the two, because they derive from different mechanisms that require the different sets of solutions.

The NGO now introduces the new policy and chooses the household randomly to provide the eligibility to purchase it. All the individuals can buy the existing policy, but only the randomly chosen eligibles can buy the new policy. We denote the utilization under new policy as y_1, the utilization under the existing policy as y_0. For simplicity, we assume that there are only compliers and noncpompliers. We will estimate the new policy's net impact on the health care utilization y over the existing policy.

III.2 Impact on Utilization

Given that there is no randomization for the existing policy, we need to focus only on the new policy. The impact, in terms of ITT, is given by:

$$ITT = \mathcal{E}[y|z = 1] - \mathcal{E}[y|z = 0],$$

where y denotes the utilization. The impact should be considered as the net of existing policy availability: the ineligibles are free to purchase the existing policy or not to.

III.3 Moral Hazard

Assume that the new policy uses an experience rating (ER) contract or bonus-malus system[*2], and other features are identical to the existing policy. Then $\mathcal{E}[y_1 - y_0]$ is the difference in utilization due to ER. In light of theory, ATE of new policy on the existing policy holders $\mathcal{E}[y_1 - y_0 | e = 1]$, where $e = 1$ denotes the indicator of existing policy holder, gives the effects of ER in reducing moral hazard. This is negative if it does reduce moral hazard. This follows since an ER policy gives an incentive to suppress utilization, through more intensive prevention efforts (curving *ex ante* moral hazard) and deterrence of overuse (curving *ex post* moral hazard). If it does make a difference in utilization, it shows that an effort can make a difference, satisfying the very definition of moral hazard. ATE on the rest of the eligibles $\mathcal{E}[y_1 - y_0 | e = 0]$,

[*2] A contract that discounts the premium if there was no claim, and increases the premium if there was a claim.

with the nonholders of existing policy denoted as $e = 0$, would give the effects of introducing the new policy on this population, which is most likely to be positive. Thus ATE of new policy $\mathcal{E}[y_1 - y_0]$, which is a weighted average of ATE on the existing policy holders and ATE on the nonholders, can either be positive or negative.

Allowing for the possibility that there may be eligible households (noncompliers) who do not prefer the new policy, we can set the bounds to ATE with:

$$ITT \leqslant ATE \leqslant \frac{1}{w_c} ITT. \quad (4)$$

w_c is the percentage of eligible households who purchase the new policy.

III.4 Adverse Selection

Suppose that the new policy is offered at the low premium but with higher deductibles (or higher copayment). The existing policy is offered at the higher premium but with lower deductibles (or lower copayment) when compared with the new policy. An ITT estimate of the new policy is $w_c\mathcal{E}[y_1 - y_0|i = c]$. ATE of new policy $\mathcal{E}[y_1 - y_0]$ is difference in utilization between the group with eligibility to buy the new policy and the group without it.

In theory, ATE of new policy on the existing policy holders $\mathcal{E}[y_1 - y_0|e = 1]$, where $e = 1$ denotes the indicator of existing policy holder, gives the effects of high deductibles in reducing adverse selection. Everything else is the same with the moral hazard example. Note, however, that they may wind up engaging in moral hazard, after they are insured. So ATE will give the aggregate effects of both adverse selection and moral hazard. To remove the effect of moral hazard from this, one needs to incorporate additional randomization (Ito and Kono, 2007). We will assume that both groups engage to the same extent in moral hazard, so its effect will be removed after we differenced ATEs of the two. Bounds are given exactly the same as in (4).

Reference

Imbens, Guido W., and Donald B. Rubin (1997), "Estimating Outcome Distributions for Compliers in Instrumental Variables Models", *Review of Economic Studies*, Vol. 64, Issue 4 (October), pp.555-574.

Ito, Seiro, and Hisaki Kono (2007), "An Empirical Assessment of the Problems in Microinsurance in India: A Conceptual Note", *mimeo*.