ESTIMATING AVERAGE TREATMENT EFFECTS: INTRODUCTION

Jeff Wooldridge
Michigan State University
BGSE/IZA Course in Microeconometrics
July 2009

1. Introduction
2. Basic Concepts

1. Introduction

• What kinds of questions can we answer using a "modern" approach to treatment effect estimation? Here are some examples:

What are the effects of a job training program on employment or labor earnings?

What are the effects of a school voucher program on student performance?

Does a certain medical intervention increase the likelihood of survival?

• The main issue in program evaluation concerns the nature of the intervention, or "treatment."

• For example, is the "treatment" randomly assigned? Hardly ever in economics, and problematical even in clinical trials because those chosen to be eligible can and do opt out. But there is a push in some fields, for example, development economics, to use more experiments.

• With retrospective or observational data, a reasonable possibility is to assume that treatment is effectively randomly assigned conditional on observable covariates. ("Unconfoundedness" or "ignorability" of treatment or "selection on observables." Sometimes called "exogenous treatment.")

• Or, does assignment depend fundamentally on unobservables, where the dependence cannot be broken by controlling for observables? ("Confounded" assignment or "selection on unobservables" or "endogenous treatment")

• Often there is a component of self-selection in program evaluation.

• Broadly speaking, approaches to treament effect estimation fall into one of three situations: (1) Assume unconfoundedness of treatment, and then worry about how to exploit it in estimation; (2) Allow self-selection on unobservables but exploit an exogenous instrumental variable; (3) Exploit a "regression discontinuity" design, where the treatment is determined (or its probability) as a discontinuous function of observed variable.

2. Basic Concepts

Counterfactual Outcomes and Parameters of Interest

● First assume a binary treatment. For each population unit, two possible outcomes: $Y(0)$ (the outcome without treatment) and $Y(1)$ (the outcome with treatment). The binary "treatment" indicator is W, where $W = 1$ denotes "treatment." The nature of $Y(0)$ and $Y(1)$ – discrete, continuous, some mix – is, for now, unspecified. (The generality this affords is one of the attractions of the **Rubin Causal Model**.)

● The gain from treatment is

$$Y(1) - Y(0). \tag{1}$$

• For a particular unit i, the gain from treatment is

$$Y_i(1) - Y_i(0).$$

If we could observe these gains for a random sample, the problem would be easy: just average the gain across the random sample.

• Problem: For each unit i, only one of $Y_i(0)$ and $Y_i(1)$ is observed.

• In effect, we have a missing data problem (even though we will eventually assume a random sample of units).

• Two parameters are of primary interest. The **average treatment effect** (**ATE**) is

$$\tau_{ate} = E[Y(1) - Y(0)]. \tag{2}$$

The expected gain for a randomly selected unit from the population. This is sometimes called the *average causal effect.*

• The **average treatment effect on the treated** (**ATT**) is the average gain from treatment for those who actually were treated:

$$\tau_{att} = E[Y(1) - Y(0)|W = 1] \tag{3}$$

• With heterogeneous treatment effects, (2) and (3) can be very different. ATE averages across gain from units that might never be subject to treatment.

• Important point: τ_{ate} and τ_{att} are defined without reference to a model or a discussion of the nature of the treatment. In particular, these definitions hold when whether assignment is randomized, unconfounded, or endogenous.

• Not suprisingly, how we estimate τ_{ate} and τ_{att} depends on what we assume about treatment assignment.

• We can also define ATEs and ATTs conditional on a set of observed covariates; in fact, some approaches to estimating τ_{ate} and τ_{att} rely on first estimating conditional average treatment effects.

Sampling Assumptions

• Assume independent, identically distributed observations from the underlying population. The data we would like to have is $\{(Y_i(0), Y_i(1)) \ : i = 1,\ldots,N\}$, but we only observe W_i and

$$Y_i = (1 - W_i)Y_i(0) + W_iY_i(1) = Y_i(0) + W_i[Y_i(1) - Y_i(0)]. \qquad (4)$$

• Random sampling rules out treatment of one unit having an effect on other units. (So the "stable unit treatment value assumption," or SUTVA, is in force.)

Estimation under Random Assignment

• Strongest form of random assignment: $[Y(0), Y(1)]$ is independent of W. Then

$$E(Y|W = 1) - E(Y|W = 0) = E[Y(1)) - E(Y(0)] = \tau_{ate} = \tau_{att} \qquad (5)$$

under mean independence and the means on the left hand side can be estimated by using sample averages on the two subsamples.

• The randomization of treatment needed for the simple comparison-of-means estimator to consistently estimate the ATE is rare in practice. Sometimes randomizaton of eligibility is reasonable, and we will consider such scenarios later.

Multiple Treatments

• If the treatment W_i takes on $G + 1$ levels, say $\{0, 1, \ldots, G\}$, it is straightforward to extend the counterfactual framework. Simply let $Y(0), \ldots, Y(G)$ denote the counterfactual outcomes associated with each level of treatment. If

$$\mu_g = E[Y(g)]$$

then we can define the expected gain in going from treatment level $g - 1$ to g as $\mu_g - \mu_{g-1}$. In some cases, $Y(0)$ might denote the response under no treatment, but generally the different values of W simply denote different treatment arms.