LECTURE NOTE 15

TREATMENT EFFECT HETEROGENEITY

1. POTENTIAL OUTCOMES AND SELECTION BIAS

As in Note 14, Section 9, there is a potential outcome function $Y_i(\cdot\,,\cdot\,)$. It can be evaluated at any level t of the treatment and level s of the subsidy. Then $Y_i(t,s)$ is a random variable, whose realized value is the outcome for i at treatment level t and subsidy level s. As t and s vary, we have a set of potential outcomes. Assume that the distribution of the potential outcome $Y_i(t,s)$ does not depend on s: for all (feasible) values of t, s_1, and s_2,

$$Y_i(t,s_1) \stackrel{d}{=} Y_i(t,s_2).$$

So we can write the potential outcome function as a function just of the treatment level:

$$Y_i(t,s) = Y_i(t).$$

Suppose that the treatment level takes on only two values, which we shall denote by $t=0$ and $t=1$. Let β_0 and β_1 denote the expected values of the two potential outcomes:

$$E[Y_i(0)] = \beta_0,$$

$$E[Y_i(1)] = \beta_1.$$

Define the prediction errors

$$U_{i0} = Y_i(0) - E[Y_i(0)],$$

$$U_{i1} = Y_i(1) - E[Y_i(1)],$$

giving two equations for the potential outcomes:

$$Y_i(0) = \beta_0 + U_{i0}, \tag{1}$$

$$Y_i(1) = \beta_1 + U_{i1}. \tag{2}$$

Let θ denote the average treatment effect:

$$\text{ATE} = \beta_1 - \beta_0 \equiv \theta.$$

Selection bias arises if the assigned treatment, T_i, is correlated with the potential outcomes. We need a flexible way to model this relationship and to allow T_i to be related to the assigned subsidy, S_i. We shall use a threshold crossing model, which compares a latent variable, V_i, to a function of the subsidy, $g(S_i)$:

$$T_i = \begin{cases} 1, & \text{if } V_i \leq g(S_i); \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Suppose that S_i consists of J variables: $S_i' = (S_{i1}, \ldots, S_{iJ})$. The function g maps a subset of $\mathcal{R}^J$ to the interval $[0, 1]$. The function g is not given; it is unknown and unrestricted except for being differentiable. Since g is unrestricted, it is not restrictive to assume that

$$V_i \,|\, S_i = s \sim \text{Uniform}[0, 1]. \tag{4}$$

Then we have

$$\text{Prob}(T_i = 1 \,|\, S_i = s) = \text{Prob}(V_i \leq g(s)) = g(s), \tag{5}$$

which is unrestricted.

The observed outcome Y_i is the potential outcome function $Y_i(\cdot)$ evaluated at the assigned (observed) treatment T_i:

$$\begin{aligned} Y_i = Y_i(T_i) &= Y_i(0) + T_i[Y_i(1) - Y_i(0)] \\ &= \beta_0 + T_i\theta + [U_{i0} + T_i(U_{i1} - U_{i0})]. \end{aligned} \tag{6}$$

The conditional expectation of observed outcome conditional on observed treatment is

$$E(Y_i \,|\, T_i) = \beta_0 + T_i\theta + E(U_{i0} \,|\, T_i) + T_i E(U_{i1} - U_{i0} \,|\, T_i).$$

The predictive effect of the treatment is

$$\begin{aligned} \text{PE} &= E(Y_i \,|\, T_i = 1) - E(Y_i \,|\, T_i = 0) \\ &= \theta + E(U_{i1} \,|\, T_i = 1) - E(U_{i0} \,|\, T_i = 0). \end{aligned} \tag{7}$$

The selection bias problem is that this predictive effect does not, in general, equal the average treatment effect if T_i is correlated with U_{i0} or with U_{i1}.

2. RANDOM ASSIGNMENT

If the treatment is randomly assigned, then it is independent of the potential outcomes:

$$\{Y_i(0), Y_i(1)\} \underline{\perp\!\!\!\perp} T_i.$$

In that case,

$$E(U_{i1} \,|\, T_i = 1) = E(U_{i1}) = 0, \quad E(U_{i0} \,|\, T_i = 0) = E(U_{i0}) = 0,$$

and the predictive effect in (7) equals the average treatment effect.

Now suppose that the treatment is not randomly assigned, but the subsidy S_i is randomly assigned:

$$\{Y_{i0}, Y_{i1}, V_i\} \underline{\perp\!\!\!\perp} S_i. \tag{8}$$

The reduced-form conditional expectations of the endogenous T and Y given the exogenous S will play a key role in the analysis.

$$E(T_i \,|\, S_i = s) = \text{Prob}(T_i = 1 \,|\, S_i = s) = g(s). \tag{9}$$

Here is the key step in obtaining $E(Y_i \,|\, S_i = s)$:

$$\begin{aligned} E[T_i(U_{i1} - U_{i0}) \,|\, S_i = s] &= E[E[T_i(U_{i1} - U_{i0}) \,|\, S_i = s, V_i] \,|\, S_i = s] \\ &= E[1(V_i \leq g(s))E(U_{i1} - U_{i0} \,|\, V_i) \,|\, S_i = s] \\ &= \int_0^1 1(v \leq g(s))E(U_{i1} - U_{i0} \,|\, V_i = v)\, dv \\ &= \int_0^{g(s)} E(U_{i1} - U_{i0} \,|\, V_i = v)\, dv. \end{aligned} \tag{10}$$

We have used

$$E(U_{i1} - U_{i0} \,|\, S_i = s, V_i) = E(U_{i1} - U_{i0} \,|\, V_i). \tag{11}$$

Because S_i is independent of $\{U_{i0}, U_{i1}, V_i\}$, it is independent of $\{U_{i0}, U_{i1}\}$ conditional on V_i, which implies (11). Now we can obtain $E(Y_i \mid S_i = s)$ from (6), using (9), (10), and $E(U_{i0} \mid S_i = s) = 0$. This gives

Claim 1. (Reduced Form)

$$E(T_i \mid S_i = s) = g(s), \tag{12}$$

$$E(Y_i \mid S_i = s) = \beta_0 + g(s)\theta + \int_0^{g(s)} E(U_{i1} - U_{i0} \mid V_i = v)\, dv. \tag{13}$$

3. MARGINAL TREATMENT EFFECT

We need to think about that latent variable V_i that is used to model treatment assignment. An important special case of our model has the assigned treatment T_i independent of the gain from treatment $(Y_i(1) - Y_i(0))$, conditional on S_i. Then V_i is independent of $(U_{i1} - U_{i0})$, and the reduced form for Y_i reduces to

$$E(Y_i \mid S_i = s) = \beta_0 + g(s)\theta. \tag{14}$$

In this case, the reduced form analysis for the (simple) IV model in Note 14, Section 10 becomes relevant:

$$E^*(T_i \mid 1, S_i) = \alpha_1 + S_i'\pi_1, \tag{15}$$

$$\begin{aligned} E^*(Y_i \mid 1, S_i) &= \beta_0 + E^*[g(S_i) \mid 1, S_i]\theta \\ &= \beta_0 + (\alpha_1 + S_i'\pi_1)\theta \\ &= (\beta_0 + \theta\alpha_1) + S_i'(\theta\pi_1) \\ &= \alpha_2 + S_i'\pi_2, \end{aligned} \tag{16}$$

where $\alpha_2 = \beta_0 + \theta\alpha_1$ and

$$\pi_2 = \theta\pi_1. \tag{17}$$

We have used

$$E^*[E(T_i \mid S_i) \mid 1, S_i] = E^*(T_i \mid 1, S_i)$$

—see Claim 1 in Section 2 of Note 2.

Now consider the general case where $E(U_{i1} - U_{i0} \,|\, V_i) \neq 0$. The *marginal treatment effect* is a function defined by

$$\text{MTE}(v) = E[Y_i(1) - Y_i(0) \,|\, V_i = v] = \theta + E(U_{i1} - U_{i0} \,|\, V_i = v).$$

The marginal treatment effect evaluated at v is the average treatment effect for the subpopulation with $V_i = v$. This subpopulation will have their treatment assignment change from 0 to 1 for a small change in the subsidy that has $g(s)$ go from a bit below v to a bit above v. The average treatment effect can be obtained by integrating the marginal treatment effect:

$$\text{ATE} = E[E(Y_i(1) - Y_i(0) \,|\, V_i)] = \int_0^1 \text{MTE}(v)\, dv.$$

In the general case, the reduced form in Claim 1 is more complex than (15) and (16). But the key insight coming out of the simple reduced form is that the average treatment effect is identified by a ratio of reduced form slopes. So we should look at a ratio of partial derivatives.

Claim 2.

$$\frac{\frac{\partial E(Y_i \,|\, S_i = s)}{\partial s_j}}{\frac{\partial E(T_i \,|\, S_i = s)}{\partial s_j}} = \text{MTE}(g(s)) \qquad (j = 1, \ldots, J).$$

Proof.

$$\frac{\partial E(Y_i \,|\, S_i = s)}{\partial s_j} = \frac{\partial g(s)}{\partial s_j} \cdot \theta + E(U_i - U_{i0} \,|\, V_i = g(s)) \cdot \frac{\partial g(s)}{\partial s_j},$$

$$\frac{\partial E(T_i \,|\, S_i = s)}{\partial s_j} = \frac{\partial g(s)}{\partial s_j}.$$

The ratio of these partial derivatives is

$$\theta + E(U_{i1} - U_{i0} \,|\, V_i = g(s)) = \text{MTE}(g(s)). \quad \diamond$$

Suppose that S consists of a single variable that has continuous variation over some interval. In order to apply Claim 2, we first need to obtain flexible approximations to the conditional expectation functions. This could be done using least-squares estimates of a linear predictor based on a polynomial (or some other series expansion) in S. With the binary variable T, we might use a polynomial inside a probit or logit function, and use a maximum likelihood estimator. Then we can take derivatives at various values for S, form ratios, and examine how the estimated marginal treatment effect varies over the interval (or intervals) where S has continuous variation.

This is not feasible if, for example, S takes on only two values. The next section applies our framework to that case.

4. LOCAL AVERAGE TREATMENT EFFECT

Suppose that the subsidy takes on only two values, which we shall denote by $s = 0$ and $s = 1$. Assume that the instrumental variable is relevant: $E(T_i \,|\, S_i = 1) \neq E(T_i \,|\, S_i = 0)$, and label the subsidy values so that $g(0) < g(1)$. Now the analog of a ratio of derivatives is the ratio of differences:

$$\begin{aligned}
\frac{E(Y_i \,|\, S_i = 1) - E(Y_i \,|\, S_i = 0)}{E(T_i \,|\, S_i = 1) - E(T_i \,|\, S_i = 0)} &= \theta + \frac{1}{g(1) - g(0)} \int_{g(0)}^{g(1)} E(U_{i1} - U_{i0} \,|\, V_i = v)\, dv \quad (18) \\
&= \theta + E[U_{i1} - U_{i0} \,|\, g(0) \leq V_i \leq g(1)] \\
&= E[Y_i(1) - Y_i(0) \,|\, g(0) \leq V_i \leq g(1)] \\
&= \text{LATE}.
\end{aligned}$$

If an individual has $g(0) < V_i \leq g(1)$, then the individual has $T_i = 1$ if $S_i = 1$ and $T_i = 0$ otherwise. If $V_i \leq g(0)$, then the individual has $T_i = 1$ regardless of the subsidy value (always taker), and if $g(1) < V_i$, then $T_i = 0$ regardless of the subsidy value (never taker). So the difference ratio in (18) identifies an average treatment effect for the compliers; this treatment effect is known as the *local average treatment effect*. It is the average treatment effect for the subpopulation for whom the subsidy has an effect.

The sample analog is the ratio of differences of sample means:

$$\frac{(\bar{Y} \mid S = 1) - (\bar{Y} \mid S = 0)}{(\bar{T} \mid S = 1) - (\bar{T} \mid S = 0)} \xrightarrow{p} \text{LATE}.$$

This estimator is known as the *Wald estimator*. Because S_i consists of a single binary variable, the Wald estimator equals the ratio $(\hat{\pi}_2/\hat{\pi}_1)$ of least-squares slope coefficients from the least-squares fits $\hat{T}_i = \hat{\alpha}_1 + S_i\hat{\pi}_1$ and $\hat{Y}_i = \hat{\alpha}_2 + S_i\hat{\pi}_2$. In this just-identified case, this equals the orthogonality condition (IV) estimator (sample $\text{Cov}(S, Y)$/sample $\text{Cov}(S, T)$), which equals the two-stage least-squares estimator.