

# 14.661: Recitation 5

Chris Walters

October 15, 2010

## 1 The Potential Outcomes Framework

### 1.1 Setup

In labor we are often interested in estimating the causal effect of some treatment on some outcome. For example, we might want to know the effect of having children on maternal labor supply, as in Angrist and Evans (1998). To think about this clearly, let's introduce some notation:

$Y_{1i}$  =  $i$ 's outcome if she is treated

$Y_{0i}$  =  $i$ 's outcome if she is not treated

In the fertility example,  $Y_{1i}$  is woman  $i$ 's LFP if she has more than 2 kids, and  $Y_{0i}$  is her LFP if she doesn't. We usually denote treatment using the notation

$$D_i = \begin{cases} 1, & i \text{ got treated} \\ 0, & i \text{ did not get treated} \end{cases}$$

In our example,  $D_i$  is an indicator for having more than 2 kids.

We are now in a position to think about causal effects. In particular, let's define

$$\beta_i \equiv Y_{1i} - Y_{0i}$$

This is the **causal effect of treatment for individual  $i$** ; it is the difference between the outcomes she would experience in the treated and non-treated states. The central problem in econometric inference is that we can never observe both  $Y_{1i}$  and  $Y_{0i}$  for a single individual. Instead, we only observe one or the other, depending on actual treatment status. That is, we observe

$$Y_i = \begin{cases} Y_{1i} & \text{if } D_i = 1 \\ Y_{0i} & \text{if } D_i = 0 \end{cases}$$

or (to write this more compactly)

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i}) D_i$$

$$Y_i = Y_{0i} + \beta_i D_i$$

For now, let's assume **constant effects**:  $\beta_i = \beta \forall i$ . Furthermore, let's define  $\alpha \equiv E[Y_{0i}]$ . Then we can write

$$Y_i = \alpha + \beta D_i + \epsilon_i \tag{1}$$

where

$$\epsilon_i = Y_{0i} - E[Y_{0i}]$$

This looks like a regression. But is it?

## 1.2 Distinction: Regressions vs. Causal Models

Above, we've written something that looks like the bivariate regressions that we are used to running. But it is important to realize that it is a **causal model** – as written,  $\beta$  is the causal effect of  $D_i$  on  $Y_i$ . We could write down a second equation:

$$Y_i = \alpha_R + \beta_R D_i + e_i \tag{2}$$

This is the **population regression of  $Y_i$  on  $D_i$** , with coefficients defined by

$$\begin{pmatrix} \alpha_R \\ \beta_R \end{pmatrix} = \arg \min_{(a,b)'} E[(Y_i - a - bD_i)^2]$$

The FOC for this problem is

$$\begin{aligned} E[D_i(Y_i - \alpha_R - \beta_R D_i)] &= 0 \\ \implies E[D_i e_i] &= 0 \end{aligned}$$

That is, the coefficients  $\alpha_R$  and  $\beta_R$  are **defined** to make  $D_i$  and  $e_i$  orthogonal.  $\alpha_R$  and  $\beta_R$  are the coefficients that we would get by running OLS in arbitrarily large samples; they are the solutions to the population least squares problem.

The crucial point to understand is that equation (1) **may not be a regression**. In that model, things were defined in terms of potential outcomes, and  $\beta$  has a causal interpretation. It is not necessary that  $D_i$  and  $\epsilon_i$  are uncorrelated. If they are correlated, regressing  $Y_i$  on  $D_i$  will produce  $\alpha_R$  and  $\beta_R$ , which are not the same as  $\alpha$  and  $\beta$ ! In contrast, the regression error  $e_i$  is uncorrelated with  $D_i$  by definition.

## 1.3 When will regression give us causal parameters?

As discussed in section 1.2, the coefficients that we get from regressing  $Y_i$  on  $D_i$  ( $\alpha_R$  and  $\beta_R$ ) may not give us the parameters of the causal relationship we are interested in ( $\alpha$  and  $\beta$ ). Since  $D_i$  is a dummy variable, we know that running a regression will produce a simple difference in means:

$$\beta_R = \frac{Cov(D_i, Y_i)}{Var(D_i)}$$

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

We can use our causal model to work this out. It says that

$$\begin{aligned}
\beta_R &= E[\alpha + \beta D_i + \epsilon_i|D_i = 1] - E[\alpha + \beta D_i + \epsilon_i|D_i = 0] \\
&= \beta + E[\epsilon_i|D_i = 1] - E[\epsilon_i|D_i = 0] \\
&= \beta + \{E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0]\}
\end{aligned}$$

That is, regression gives us the causal parameter of interest, **plus the difference in average  $Y_0$  between treated and non-treated individuals**. This may not be zero. In the fertility example, if there is any reason that potential outcomes differ systematically between women who do and don't have a third child (for example, if women who have three kids would have earned less than women who have two anyway), regression will not produce the causal  $\beta$ . A necessary and sufficient condition for  $\beta_R = \beta$  is then

$$Y_{0i} \perp D_i$$

## 2 IV in the Potential Outcomes Framework

In the fertility example (and many others), it is implausible that  $Y_{0i}$  is unrelated to  $D_i$ , so the observed relationship between  $Y$  and  $D$  will fail to give us the causal parameter of interest. In situations like this, we need an instrument. Let's suppose we're considering  $Z_i \in \{0, 1\}$ . The first stage and exclusion restrictions are:

1.  $E[D_i|Z_i = 1] \neq E[D_i|Z_i = 0]$
2.  $E[Y_{0i}|Z_i = 1] = E[Y_{0i}|Z_i = 0]$

This tells us that  $Z_i$  shifts  $D_i$  on average, but  $Y_{0i}$  is orthogonal to  $Z_i$ . IV will produce

$$\begin{aligned}
\beta_{IV} &= \frac{Cov(Y_i, Z_i)}{Cov(D_i, Z_i)} \\
&= \frac{Cov(Y_i, Z_i)/Var(Z_i)}{Cov(D_i, Z_i)/Var(Z_i)} \\
&= \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]}
\end{aligned}$$

This is the population Wald formula (or the Wald estimand). It is the ratio of the reduced form (the effect of  $Z$  on  $Y$ ) to the first stage (the effect of  $Z$  on  $D$ ). Under assumptions (1) and (2), this will give us

$$\begin{aligned}\beta_{IV} &= \frac{E[\alpha + \beta D_i + \epsilon_i | Z_i = 1] - E[\alpha + \beta D_i + \epsilon_i | Z_i = 0]}{E[D_i | Z_i = 1] - E[D_i | Z_i = 0]} \\ &= \frac{\beta (E[D_i | Z_i = 1] - E[D_i | Z_i = 0]) + (E[Y_{0i} | Z_i = 1] - E[Y_{0i} | Z_i = 0])}{E[D_i | Z_i = 1] - E[D_i | Z_i = 0]} \\ &= \beta\end{aligned}$$

So IV yields the causal parameter of interest.

Angrist and Evans use two instruments to estimate the effect of family size on maternal LFP: Twinning, and sibling sex composition. To see the motivation for the twins instrument, think about limiting your sample to women who have at least two kids. In this sample, there will be some women who would like to stop at 2. But some of these women will have twins at second birth, so they will end up with 3. We should therefore expect the “twins at second birth” group to have more children on average, so there is a first stage. Furthermore, since having twins rather than a single at your second birth is essentially random (subject to the age and ethnicity caveats Josh gave in class), we might think that it’s uncorrelated with anything else that affects labor force participation; this is the exclusion restriction.

To see the case for the sex composition instrument, suppose that some families have a preference for a “diversified sibling sex portfolio.” They would like to have kids of both sexes. Consider the sample of families who have decided to have a second child. In some of these families, the second birth will result in a boy/girl combo; in the rest, it will not. Some of the families who end up with two kids of the same sex will choose to keep trying and have a third in order to achieve a diversified portfolio. Therefore, families who end up with a matched pair will have more children on average than families who get diversification at second birth, and we have a first stage. Since the sex of the second child is random and seems unlikely to affect LFP through other channels, we might expect the exclusion restriction to hold.

Note that the exclusion restriction is subtle. In particular, **random assignment of  $Z$  does not guarantee the exclusion restriction**. We know that random assignment of  $Z$  will give us a causal interpretation of the reduced form, so we know that getting twins at second birth decreases maternal LFP (assuming twinning is truly random). But the exclusion restriction says that family size is the **only channel** through which maternal LFP is affected by twinning. To see how this might fail even with random assignment, suppose that a twin birth is physically more difficult for the mother, and therefore more likely to lead to health complications for her. Mothers who have twins might therefore work less due to poor health rather than because of family size. We might think of this as a direct effect of twinning on  $Y_{0i}$ , which is a little awkward in the framework that we’ve written down. This particular channel seems unlikely, but the point is that one must always make a case for the exclusion restriction, and you can’t just appeal to random assignment!

### 3 Heterogeneous Treatment Effects

#### 3.1 Setup

Now let’s relax the clearly unrealistic assumption that treatment effects are the same for everyone. Our causal model is then

$$Y_i = \alpha + \beta_i D_i + \epsilon_i$$

In this case, what we want to estimate is no longer clear. A couple of plausible candidates are:

$$ATE : E[\beta_i]$$

$$TOT : E[\beta_i | D_i = 1]$$

When will regression give us something interesting? As before, the population regression coefficient is

$$\begin{aligned} \beta_R &= E[Y_i | D_i = 1] - E[Y_i | D_i = 0] \\ &= E[\alpha + \beta_i D_i + \epsilon_i | D_i = 1] - E[\alpha + \beta_i D_i + \epsilon_i | D_i = 0] \\ &= E[\beta_i | D_i = 1] + E[\epsilon_i | D_i = 1] - E[\epsilon_i | D_i = 0] \\ &= E[\beta_i | D_i = 1] + \{E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]\} \end{aligned}$$

So as before, regression gives us an object of interest (the TOT) plus the difference in average  $Y_0$  between those who take the treatment and those who don't.

### 3.2 IV with heterogeneous treatment effects

Suppose we are not confident that  $D_i$  is orthogonal to  $Y_{0i}$ , so we want to do IV. We need to add more notation. Let  $Z_i \in \{0, 1\}$  be our candidate for an instrument, and  $Y_i(d, z)$  be  $i$ 's potential outcome when  $D_i = d$  and  $Z_i = z$ . We can also define treatment status as a function of the instrument:  $D_{1i}$  is treatment status if  $Z_i = 1$  and likewise for  $D_{0i}$ , so

$$D_i = D_{0i} + (D_{1i} - D_{0i}) Z_i$$

1. Independence:  $Y_i(d, z)$ ,  $D_{0i}$ ,  $D_{1i}$  are independent of  $Z_i$
2. Exclusion:  $Y_i(d, 0) = Y_i(d, 1) \equiv Y_{di} \forall d$
3. First stage:  $E[D_{1i} - D_{0i}] \neq 0$
4. Monotonicity:  $D_{1i} \geq D_{0i} \forall i$

What we were calling the "exclusion restriction" before actually bundled requirements (1) and (2). Now, when we do IV, we obtain

$$\begin{aligned} \beta_{IV} &= \frac{E[Y_i | Z_i = 1] - E[Y_i | Z_i = 0]}{E[D_i | Z_i = 1] - E[D_i | Z_i = 0]} \\ &= \frac{E[Y_{0i} + \beta_i D_i | Z_i = 1] - E[Y_{0i} + \beta_i D_i | Z_i = 0]}{E[D_i | Z_i = 1] - E[D_i | Z_i = 0]} \\ &= \frac{E[Y_{0i} + \beta_i D_{1i} | Z_i = 1] - E[Y_{0i} + \beta_i D_{0i} | Z_i = 0]}{E[D_{1i} | Z_i = 1] - E[D_{0i} | Z_i = 0]} \end{aligned}$$

Now, using independence, we have

$$= \frac{E[Y_{0i} + \beta_i D_{1i}] - E[Y_{0i} + \beta_i D_{0i}]}{E[D_{1i}] - E[D_{0i}]}$$

$$\begin{aligned}
&= \frac{E[\beta_i D_{1i}] - E[\beta_i D_{0i}]}{E[D_{1i}] - E[D_{0i}]} \\
&= \frac{E[\beta_i (D_{1i} - D_{0i})]}{E[D_{1i} - D_{0i}]} \\
&= \frac{E[\beta_i | D_{1i} > D_{0i}] Pr[D_{1i} > D_{0i}] + 0 Pr[D_{1i} = D_{0i}]}{Pr[D_{1i} > D_{0i}]} \\
&= E[\beta_i | D_{1i} > D_{0i}]
\end{aligned}$$

This is the Local Average Treatment Effect, or LATE. It tells us that IV estimates the average treatment effect FOR INDIVIDUALS WHOSE TREATMENT STATUS IS SHIFTED BY THE INSTRUMENT. We can think about 4 possible groups:

1. Always takers:  $D_{1i} = D_{0i} = 1$
2. Never takers:  $D_{1i} = D_{0i} = 0$
3. Compliers:  $D_{1i} = 1, D_{0i} = 0$
4. Defiers:  $D_{1i} = 0, D_{0i} = 1$

Monotonicity rules out defiers, and LATE is the average causal effect on compliers.

MIT OpenCourseWare  
<http://ocw.mit.edu>

14.661 Labor Economics I  
Fall 2010

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.