# Manski Bounds

Alex Tabarrok George Mason Universtiy

January 2008

## 1 Manski Bounds

- Let $t = 1$ denote residential treatment, $t = 0$ denotes nonresidential or non treatment.
- The measured outcome, recidivism, is denoted $y, y = 1$ is recidivate, $y = 0$ is does not recidivate. Where recidivate is re-offending within a 2 year period.
- $y(t) = 1$ means the individual would recidivate if assigned treatment $t$ and $y(t) = 0$ means the individual would not recidivate given treatment $t$. Note that many of these outcomes will be counter-factual.
- Let $z = 1$ if an offender is actually sentenced to residential treatment and $z = 0$ if assigned to non treatment then $P[z = 1] =$ prob of treatment and $P[z = 0] =$ prob of nontreatment. I will simplify this notation to $P_z$ and $(1 - P_z)$.

Now remember the ATE or what Manski and Nagin call the classical treatment effect

$$CTE = P[y(1) = 1] - P[y(0) = 1]$$

this is simply the expected outcome if everyone were assigned to treatment minus the expected outcome if everyone were assigned to nontreatment.

Now let's expand each element of the CTE

$$P[y(1) = 1] = P[y(1) = 1|z = 1] * P_z + P[y(1) = 1|z = 0] * (1 - P_z)$$

this is just a definition or a use of the law of total probability. In other words, since either $z = 1$ or $z = 0$ then

$$P[anything] = P[anything|z = 1) * P_z + P[anything|z = 0] * (1 - P_z)$$

Similarly,

$$\begin{aligned} P[y(0) &= 1] = P[y(0) = 1|z = 1] * P[z = 1] + \\ P[y(0) &= 1|z = 0] * P[z = 0] \end{aligned}$$

So we can also write the CTE as

$$\begin{aligned} CTE &= (P[y(1) = 1|z = 1] * P_z + P[y(1) = 1|z = 0] * (1 - P_z)) \\ &\quad - (P[y(0) = 1|z = 1] * P_z + P[y(0) = 1|z = 0] * (1 - P_z)) \end{aligned}$$

Now of this CTE what does the data tell us?
The data tell us:

- $P_z$
- $P[y(1) = 1|z = 1]$ = the probability of recidivism given treatment for the group actually treated.
- $P[y(0) = 1|z = 0]$ = prob of recidivism given nontreatment for the group not treated.

The data do *not* tell us::

- $P[y(1) = 1|z = 0]$ = prob. of recidivism given treatment for the group not treated. (counter-factual)
- $p[y(0) = 1|z = 1]$ = prob of recidivism given non treatment for the group treated. (counter-factual)

thus here is the CTE again with the unknown elements **bolded**:

$$\begin{aligned} CTE &= (P[y(1) = 1|z = 1] * P_z + \mathbf{P[y(1) = 1|z = 0]} * (1 - P_z)) \\ &\quad - (\mathbf{P[y(0) = 1|z = 1]} * P_z + P[y(0) = 1|z = 0] * (1 - P_z)) \end{aligned}$$

- Traditional analysis says under random assignment:

$$\mathbf{P}[\mathbf{y}(\mathbf{1}) = \mathbf{1}|\mathbf{z} = \mathbf{0}] = P[y(1) = 1|z = 1] \tag{1}$$

$$\mathbf{P}[\mathbf{y}(\mathbf{0}) = \mathbf{1}|\mathbf{z} = \mathbf{1}] = P[y(0) = 1|z = 0] \tag{2}$$

- So create a research design that makes it plausible for use to substitute these elements and identify the CTE.

- Manski says let's not assume so much about the counterfactual probabilities. What is the minimum that we can assume about these probabilities?

So what is the least that we can assume about say

$$\mathbf{P}[\mathbf{y(1)} = \mathbf{1}|\mathbf{z} = \mathbf{0}]$$

Well we can certainly assume that the lowest this probability could be is zero and the highest it could be is 1, that is we could assume that none of the non-treated group would recidivate if treated or all of them would. Similarly, with

$$\mathbf{P}[\mathbf{y(0)} = \mathbf{1}|\mathbf{z} = \mathbf{1}]$$

we can be sure that this probability is no smaller than zero and no bigger than 1, ie. if they were not treated none of the treated group would recidivate or all of them would.

Now return to the CTE, what is the lower bound or best case given that we want low recidivism scenario for treatment? It is:

$$\begin{aligned} \mathbf{P}[\mathbf{y(1)}{=}\ \mathbf{1}|\mathbf{z} = \mathbf{0}] &= \mathbf{0} \\ \mathbf{P}[\mathbf{y(0)}{=}\ \mathbf{1}|\mathbf{z} = \mathbf{1}] &= \mathbf{1} \end{aligned} \tag{Lower Bound}$$

i.e. the best case for treatment is that if the nontreated had been treated then none would have recidivated and if the treated had not been treated than all would have recidivated.

What about the upper bound (worst case - given that we want low recidivism)

$$\begin{aligned} \mathbf{P}[\mathbf{y(1)}{=}\ \mathbf{1}|\mathbf{z} = \mathbf{0}] &= \mathbf{1} \\ \mathbf{P}[\mathbf{y(0)}{=}\ \mathbf{1}|\mathbf{z} = \mathbf{1}] &= \mathbf{0} \end{aligned} \tag{Upper Bound}$$

i.e. the worst case for treatment is that if the nontreated had been treated then all of them would have recidivated and if the treated had not been treated the none would have recidivated.

Recall the CTE

$$
\begin{aligned}
CTE &= (P[y(1)=1|z=1]*P_z+\mathbf{P[y(1)=1|z=0]}*(1-P_z)) \\
&\quad -(\mathbf{P[y(0)=1|z=1]}*P_z+P[y(0)=1|z=0]*(1-P_z))
\end{aligned}
$$

Now under the best case and worst case for treatment the CTE is (substituting the Lower Bound and Upper Bound from above):

$$
\begin{aligned}
CTE_{LB} &= (P[y(1)=1|z=1]*P_z)- \\
&\quad (P_z+P[y(0)=1|z=0]*(1-P_z)) \\
CTE_{UB} &= (P[y(1)=1|z=1]*P_z+(1-P_z))- \\
&\quad (P[y(0)=1|z=0]*(1-P_z))
\end{aligned}
$$

Or putting this all together to get the *no-assumptions CTE*.

$$
\begin{aligned}
&(P[y(1)=1|z=1]*P_z)-(P[y(0)=1|z=0]*(1-P_z)+P_z)\leq \\
&CTE \\
\leq\ &(P[y(1)=1|z=1]*P_z+(1-P_z))-(P[y(0)=1|z=0]*(1-P_z))
\end{aligned}
$$

Note that there are some common elements on both sides

$$stuff-P_z\leq CTE\leq stuff+(1-P_z)$$

So we know that the no-assumptions bound on the CTE has width 1.

Since the CTE can be at most 1 and the no-assumptions bound has width 1 then 0 is always included within the possibilities of the no-assumptions bound so in one sense this is dissapointing. Realistically the no assumptions bound can't even tell us the sign of the CTE. On the other hand without any data at all the CTE can be between -1 and 1 so the data half the bound.

## 2 Data

Manski and Nagin have data on 13,197 juveniles from 1970-1974 in Utah. The data says the following

- $P_z = 0.11$
- $P[y(1) = 1|z = 1] = 0.77$ the probability of recidivism given treatment for the group actually treated.
- $P[y(0) = 1|z = 0] = 0.59$ prob of recidivism given nontreatment for the group not treated.

Note that the probability of recidivism is higher for the treated - this is actually a pretty common finding. It could be selection, of course, but we should not rule out brutalization or labeling - calling someone a juvenile delinquent may get them to act more like a juvenile delinquent or peer effects.

- The CTE assuming random assignment or what you might call the naive CTE is simply: $0.77 - 0.59 = 0.18$,treatment causes recidivism.
- The no-assumptions bound on the CTE is $-0.56 \leq CTE \leq 0.44$

thus the no-assumptions bound tells us that without assumptions the data is just not very informative - the treatment effect could reduce recidivism by about 50% or increase it by about 50%.

We can also look at the no-assumptions bound CTE in various subgroups. Manski and Nagin look at bound depending on prior referrrals to the juvenile justice system and find:

| No Assumptions Bound by Priors | | | | |
|---|---|---|---|---|
| Priors | Sample Size | Random (Naive) CTE | No assum. LB | No assum. UB |
| 0 | 7406 | 0.09 | -0.48 | .52 |
| 1 | 2719 | 0.07 | -0.65 | 0.35 |
| 2+ | 3072 | 0.02 | -0.65 | 0.35 |

# 3 Adding Assumptions

**The Outcome Optimization Model** Assume that judges estimate recidivism probabilities and treatment effects and assign offenders to the treatment that is most likely to reduce recidivism.

The outcome optimization model implies (assuming rational expectations):

$$\begin{aligned}
P[y(1) &= 1|z=1] \leq \mathbf{P[y(0) = 1|z = 1]} \\
P[y(0) &= 1|z=0] \leq \mathbf{P[y(1) = 1|z = 0]}
\end{aligned}$$

These conditions let us create new LB and UB on the CTE. Recall the CTE

$$\begin{aligned}
CTE &= (P[y(1)=1|z=1]*P_z + \mathbf{P[y(1)=1|z=0]}*(1-P_z)) \\
&\quad -(\mathbf{P[y(0)=1|z=1]}*P_z + P[y(0)=1|z=0]*(1-P_z))
\end{aligned}$$

now consider the first bolded element $\mathbf{P[y(1) = 1|z = 0]}$ under outcome optimization the least this can be is $P[y(0) = 1|z = 0]$ and for $\mathbf{P[y(0) = 1|z = 1]}$ the largest this can be is still 1. Thus the new LB or best case scenario is:

$$\begin{aligned}
CTE_{LB} &= (P[y(1)=1|z=1]*P_z + P[y(0)=1|z=0]*(1-P_z)) - \\
&\quad (P_z + P[y(0)=1|z=0]*(1-P_z)) \\
&= P[y(1)=1|z=1]*P_z - P_z = -P_z(1-P[y(1)=1|z=1]) \\
&= -P_z * P[y(1)=0|z=1]
\end{aligned}$$

what about the UB? $\mathbf{P[y(1) = 1|z = 0]}$ can be at most 1 and $\mathbf{P[y(0) = 1|z = 1]}$ is least $P[y(1) = 1|z = 1]$ making the substitutions we have:

$$\begin{aligned}
CTE &= (P[y(1)=1|z=1]*P_z + (1-P_z)) - \\
&\quad (P[y(1)=1|z=1]*P_z + P[y(0)=1|z=0]*(1-P_z)) \\
&= ((1-P_z)) - P[y(0)=1|z=0]*(1-P_z) \\
&= (1-P_z)(1-P[y(0)=1|z=0]) \\
&= (1-P_z)P[y(0)=0|z=0]
\end{aligned}$$

so putting this alltogether we can get the new CTE bounds assuming outcome maximization:

$$-P[y(1) = 0|z = 1] * P_z \leq CTE \leq P[y(0) = 0|z = 0] * (1 - P_z)$$

Using this new bound Manski and Nagin find:

| Bounds With Outcome Opt. Assum. | | |
|---|---|---|
| Priors | Outcome Opt. LB | Outcome opt. UB |
| 0 | -0.02 | .50 |
| 1 | -0.03 | 0.26 |
| 2+ | -0.04 | 0.13 |

- With outcome optimization the data is consistent with at best a two to four percent reduction in recidivism - and are also consistent with a large increase in recidivism.

- Not surprisingly if you assume that judges are doing their best to minimize recidivism then when the observed data tell you that recidivism increases with treatment the treatment effect must be small at best.

**The Skimming Model** Assume that judges assign the worst offenders to residential treatment regardless. (Perhaps because they want to punish or not look soft or protect themselves if recidivism occurs.)

Under the skimming model bounds they find:

| Bounds with Skimming Assum. | | |
|---|---|---|
| Priors | Skimming LB | Skimming UB |
| 0 | -0.48 | .09 |
| 1 | -0.65 | 0.07 |
| 2+ | -0.65 | 0.02 |

- The skimming model is consistent with very large reductions in recidivism or with small increases in recidivism.

# 4 IV plus Bounds

- Suppose that $x$ does not influence treatment effectiveness - this is called an exclusion restriction because such a variable can be excluded from the determination of treatment effectiveness but suppose that $x$ does influence treatment assignment. i.e. suppose $x$ is an IV.
- Create strata of $x$ - and find the bounds in each strata of $x$. Since $x$ doesn't influence treatment effectivness the lowest of the upper bounds in the strata is the upper bound on the effect and similarly the highest of the lower bounds in the strata is the lower bound on the effect.
- In other words, you can take the intersection of the bounds.

| Bounds on CTE using District as Exclusion Restriction and Different Assum. | | | | |
|---|---|---|---|---|
| Priors | Outcome Opt. LB | Outcome opt. UB | Skimming LB | Skimming UB |
| 0 | 0.06 | .46 | -.43 | 0.03 |
| 1 | 0.07 | 0.19 | -.59 | -0.08 |
| 2+ | 0.02 | 0.06 | -.62 | -0.11 |

So what this says is that with the outcome optimization model the district exclusion restriction lets us reject any good effect from treatment. With the skimming model, however, the district exclusion lets us reject almost any bad effect from treatment!

Hard to say what the policy conclusion is here except the following - the conclusions follow much more from the assumptions made about treatment assignment than they follow from the data. This alone is useful to know. Of course, it is also a bit depression. But although it would be nice if the data told us more but there is no use making claims that we cannot back up.