Manski Bounds

Erin Hartman

PS239 Sp2010

Feb. 17, 2010

Manski Bounds

- A non-parametric way to bound treatment effects in the face of selection bias
- Selection bias can be non-response, self-selection of treatment, unobservable outcomes (such as when a candidate does not run again)
- Manski bounds are <u>sharp</u> bounds, i.e. nothing else can be learned in face of the censored data
- The selection problem is an identification problem; it limits the internal validity of the the treatment effect.

Difference between Bound and Confidence Interval

- A bound on a treatment effect is a population concept.
- A bound expresses what could be learned about the treatment effect if one knew the treatment effect for the compliers and the probability of complying.
- A confidence interval is a sample concept.
- A confidence interval expresses the precision with which the bound is estimated when estimates of the above treatment effects are obtained from a sample of fixed size.
- A confidence interval is typically wider than the bound, but narrows to match the bound as the sample size increases.

The Perry Preschool Project

- Beginning in 1962, the Perry Preschool Project provided intensive educational and social services to a random sample of disadvantaged black children in Ypsilanti, Michigan
- Also drew a random sample as control who received no help.
- It was found that 67% of the experimental group and 49% of the control group were high-school graduates by age 19.
- Highly cited as evidence that early childhood intervention is important for children from disadvantaged backgrounds.

Let each individual i be characterized by:

$$(y_1, y_0, x, z)$$

where y_1 and y_0 are potential outcomes, x is a set of observable characteristics, and z is the treatment received.

We will be bounding the value $P(y_m \in B|x)$, the probability that the realized outcome falls in a specified set B, conditional on x.

- We are trying to determine the treatment policies that minimize and maximize $P(y_m \in B|x)$
- If y_1 and y_0 both fall in B, then y_m must be in B. Similar argument if y_0 and y_1 do not fall in B.
- Therefore, the treatment policy is relevant in those cases in which one of the two outcomes falls in B and the other does not.

No Prior Information Case

- Treatment policy minimizes if it always selects the treatment yielding the outcome not in B, i.e.:

$$y_1 \notin B \cap y_0 \in B \Rightarrow z_m = 1]$$

$$y_1 \in B \cap y_0 \notin B \Rightarrow z_m = 0]$$

- Therefore the smallest possible value of $P(y_m \in B|x)$ is

$$P(y_1 \in B \cap y_0 \in B|x)$$

- It follows that the maximum value is obtained if the treatment policy always selects the treatment yielding the outcome that is in B
- A similar argument to above gives us that the largest possible value of $P(y_m \in B|x)$ is

$$P(y_1 \in B \cup y_0 \in B|x)$$

No Prior Information Case

Therefore, the sharp bounds should be:

$$P(y_1 \in B \cap y_0 \in B|x) \leq P(y_m \in B|x) \leq P(y_1 \in B \cup y_0 \in B|x)$$

however, we only know $P(y_1|x)$ and $P(y_0|x)$, so the best computable bounds are those that are consistent with these probabilities and the bounds shown above, i.e. the largest and smallest feasible values.

In light of this, we use the Frechet computable sharp bounds:

Lower:

$$max[0, P(y_1 \in B|x) + P(y_0 \in B|x) - 1] \leq P(y_1 \in B \cap y_0 \in B|x) \leq$$

$$min[P(y_1 \in B|x), P(y_0 \in B|x)]$$

This implies that $max[0, P(y_1 \in B|x) + P(y_0 \in B|x) - 1]$ is the best computable lower bound.

No Prior Information Case

Upper: Note that

$$P(y_1 \in B \cup y_0 \in B|x) = P(y_1 \in B|x) + P(y_0 \in B|x) - P(y_1 \in B \cap y_0 \in B|x)$$

Then using Frechet on $max[0, P(y_1 \in B|x) + P(y_0 \in B|x) - 1]$ implies that

$$P(y_1 \in B \cup y_0 \in B|x) \leq min[P(y_1 \in B|x) + P(y_0 \in B|x), 1]$$

Which gives us our computable bounds:

$$max[0, P(y_1 \in B|x) + P(y_0 \in B|x) - 1] \leq P(y_m \in B|x)$$
$$\leq min[P(y_1 \in B|x) + P(y_0 \in B|x), 1]$$

PPP No Prior Information

We know:

$$P(y_1 = 1|x) = 0.67$$

$$P(y_0 = 1|x) = 0.49$$

So, our bounds, with no prior information (or assumptions), are:

$$max[0, P(y_1 \in 1|x) + P(y_0 \in 1|x) - 1] \leq P(y_m \in 1|x)$$
$$\leq min[P(y_1 \in 1|x) + P(y_0 \in 1|x), 1]$$

$$\Rightarrow \big[max[0, 0.67 + 0.49 - 1 = 0.16], min[0.67 + 0.49 = 1.16, 1]\big]$$
$$\Rightarrow \big[0.16, 1\big]$$

Ordered Outcomes

- We can tighten these bounds with some assumptions. One assumption is the monotone treatment assumption.
- This assumption says that the outcome under treatment is at least as good as the outcome under control (i.e. treatment doesn't hurt).
- In the PPP this means that getting extra social services and education does not cause a student to be less likely to graduate from high school.
- Mathematically this is:

$$y_1 \geq y_0$$

$$y_1 \in B \Rightarrow y_0 \in B$$

This tightens our bounds to:

$$P(y_1 \in B|x) \leq P(y_m \in 1|x) \leq P(y_0 \in B|x)$$

PPP Ordered Outcomes

We know:

$$P(y_1 = 1|x) = 0.67$$

$$P(y_0 = 1|x) = 0.49$$

So, our bounds using the monotonicity assumption are:

$$P(y_0 \in 1|x) \leq P(y_m \in 1|x) \leq P(y_1 \in 1|x)$$

$$\Rightarrow \left[0.49, 0.67\right]$$

Optimizing Treatment

Now we assume that the student chooses the treatment yielding the larger outcome:

$$y_m = max(y_1, y_0)$$

Or, in other words, those that chose to stick with the treatment regime were more likely to have better outcomes anyway.
This yields the bounds:

$$max[0, P(y_1 \leq t|x) + P(y_0 \leq t|x) - 1] \leq P(y_m \leq t|x)$$

$$\leq min[P(y_1 \leq t|x), P(y_0 \leq t|x)]$$

Incumbency Advantage

Say we want to know the incumbency advantage for those who narrowly win versus those who narrowly lose their first election. We only observe outcomes for those candidates who run again.

- Let Y_1 denote the potential incumbency advantage and Y_0 denote the potential outcome if the incumbent loses. Let $Z = 0$ denote that the candidate decides not to run again and $Z = 1$ denote that the candidate does.
- let π refer to the number of treated units and λ refer to the number of candidates who run again.
- In the face of no selection, what is the average treatment effect?

Incumbency Advantage

Say we want to know the incumbency advantage for those who narrowly win versus those who narrowly lose their first election. We only observe outcomes for those candidates who run again.

- Let Y_1 denote the potential incumbency advantage and Y_0 denote the potential outcome if the incumbent loses. Let $Z = 0$ denote that the candidate decides not to run again and $Z = 1$ denote that the candidate does.
- let π refer to the number of treated units and λ refer to the number of candidates who run again.
- In the face of no selection, what is the average treatment effect?

$$E[\delta] = \left\{\pi E[Y_1|T=1] + (1-\pi)E[Y_1|T=0]\right\}$$
$$-\left\{\pi E[Y_0|T=1] + (1-\pi)E[Y_0|T=0]\right\}$$

Incumbency Advantage

In the face of selection, how might we decompose the treatment effect?

Incumbency Advantage

In the face of selection, how might we decompose the treatment effect?

$$E[\delta] = \Big\{\pi(\lambda E[Y_1|T=1,Z=1] + (1-\lambda)E[Y_1|T=1,Z=0]) +$$

$$(1-\pi)(\lambda E[Y_1|T=0,Z=1] + (1-\lambda)E[Y_1|T=0,Z=0])\Big\}$$

$$-\Big\{\pi(\lambda E[Y_0|T=1,Z=1] + (1-\lambda)E[Y_0|T=1,Z=0]) +$$

$$(1-\pi)(\lambda E[Y_0|T=0,Z=1] + (1-\lambda)E[Y_0|T=0,Z=0])\Big\}$$

Incumbency Advantage

In the face of selection, how might we decompose the treatment effect?

$$E[\delta] = \Big\{\pi(\lambda E[Y_1|T=1,Z=1] + (1-\lambda)E[Y_1|T=1,Z=0]) +$$

$$(1-\pi)(\lambda E[Y_1|T=0,Z=1] + (1-\lambda)E[Y_1|T=0,Z=0])\Big\}$$

$$-\Big\{\pi(\lambda E[Y_0|T=1,Z=1] + (1-\lambda)E[Y_0|T=1,Z=0]) +$$

$$(1-\pi)(\lambda E[Y_0|T=0,Z=1] + (1-\lambda)E[Y_0|T=0,Z=0])\Big\}$$

How can we simplify this?

Incumbency Advantage

In the face of selection, how might we decompose the treatment effect?

$$E[\delta] = \Big\{\pi(\lambda E[Y_1|T=1, Z=1] + (1-\lambda)E[Y_1|T=1, Z=0]) +$$

$$(1-\pi)(\lambda E[Y_1|T=0, Z=1] + (1-\lambda)E[Y_1|T=0, Z=0])\Big\}$$

$$-\Big\{\pi(\lambda E[Y_0|T=1, Z=1] + (1-\lambda)E[Y_0|T=1, Z=0]) +$$

$$(1-\pi)(\lambda E[Y_0|T=0, Z=1] + (1-\lambda)E[Y_0|T=0, Z=0])\Big\}$$

How can we simplify this?

Assume $(Y_1, Y_0) \perp\!\!\!\perp T$

Incumbency Advantage

$$E[\delta] = \Big\{(\lambda E[Y_1|Z=1] + (1-\lambda)E[Y_1|Z=0]\Big\}$$
$$-\Big\{(\lambda E[Y_0|Z=1] + (1-\lambda)E[Y_0|Z=0]\Big\}$$

Incumbency Advantage

$$E[\delta] = \Big\{(\lambda E[Y_1|Z=1] + (1-\lambda)E[Y_1|Z=0]\Big\}$$
$$-\Big\{(\lambda E[Y_0|Z=1] + (1-\lambda)E[Y_0|Z=0]\Big\}$$

What can we observe?

Incumbency Advantage

$$E[\delta] = \Big\{(\lambda E[Y_1|Z=1] + (1-\lambda)E[Y_1|Z=0]\Big\}$$
$$-\Big\{(\lambda E[Y_0|Z=1] + (1-\lambda)E[Y_0|Z=0]\Big\}$$

What can we observe?

Then, if we estimate our treatment effect based on our observed data as:

$$\hat{\delta} = E[Y_1|T=1, Z=1] - E[Y_0|T=0, Z=1]$$

We know it is biased. Can we bound the bias?

Incumbency Advantage

Assume that the incumbency advantage is a binary outcome, bounded as [0, 1]. Then we know the treatment effect must be between [-1, 1]. We know that our estimate is off by:

$$E[\delta] = \hat{\delta} + (1 - \lambda)\big(\hat{\delta} - \{E[Y_1|Z = 0] - E[Y_0|Z = 0]\}$$

How can we bound this using the knowledge that the dependent variable is bounded?

Incumbency Advantage

Assume that the incumbency advantage is a binary outcome, bounded as [0, 1]. Then we know the treatment effect must be between [-1, 1]. We know that our estimate is off by:

$$E[\delta] = \hat{\delta} + (1-\lambda)\big(\hat{\delta} - \{E[Y_1|Z=0] - E[Y_0|Z=0]\}$$

How can we bound this using the knowledge that the dependent variable is bounded?

$$E[\delta] \in \left\{\hat{\delta} + (1-\lambda)\big(\hat{\delta} - 1\big), \hat{\delta} + (1-\lambda)\big(\hat{\delta} + 1\big)\right\}$$

Incumbency Advantage

What if we impose the monotonicity assumption? In this case we assume that $Y_1 \geq Y_0$.

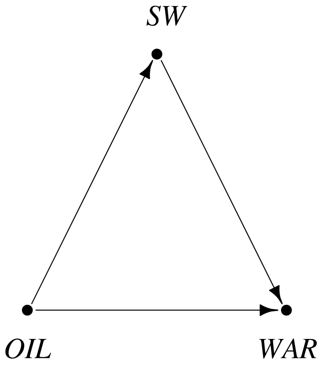
Incumbency Advantage

What if we impose the monotonicity assumption? In this case we assume that $Y_1 \geq Y_0$.

$$E[\delta] \in \left\{\hat{\delta} + (1-\lambda)\big(\hat{\delta} - 1\big), \hat{\delta}\right\}$$

Oil, Civil War, and State Weakness

What if we want to bound a mechanism specific effect? Say we believe that having oil causes civil war through the mechanism of state weakness.



Oil, Civil War, and State Weakness

Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16	1	0	0			
3	1	0	1			
9	1	1	0			
0	1	1	1			

Oil, Civil War, and State Weakness

Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16	1	0	0	{-1, 0}	{0, 1}	
3	1	0	1	{-1, 0}	{-1, 0}	
9	1	1	0	{0, 1}	{-1, 0}	
0	1	1	1	{0, 1}	{0, 1}	

Oil, Civil War, and State Weakness

Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16	1	0	0	{-1, 0}	{0, 1}	{-1, 0}
3	1	0	1	{-1, 0}	{-1, 0}	{0, 1}
9	1	1	0	{0, 1}	{-1, 0}	{-1, 0}
0	1	1	1	{0, 1}	{0, 1}	{ 0, 1}

Oil, Civil War, and State Weakness

Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16	1	0	0	{-1, 0}	{0, 1}	{-1, 0}
3	1	0	1	{-1, 0}	{-1, 0}	{0, 1}
9	1	1	0	{0, 1}	{-1, 0}	{-1, 0}
0	1	1	1	{0, 1}	{0, 1}	{ 0, 1}
						AVG = [-25/28, 3/28]

Oil, Civil War, and State Weakness

What if we use the monotone treatment assumption on the second stage?
(i.e. Assume that state weakness cannot prevent war)
Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16		0	0			
3		0	1			
9		1	0			
0		1	1			

Oil, Civil War, and State Weakness

What if we use the monotone treatment assumption on the second stage?
(i.e. Assume that state weakness cannot prevent war)
Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16		0	0	{-1, 0}	{0, 1}	
3		0	1	{-1, 0}	{$\cancel{-1}$, 0}	
9		1	0	{0, 1}	{$\cancel{-1}$, 0}	
0		1	1	{0, 1}	{0, 1}	

Oil, Civil War, and State Weakness

What if we use the monotone treatment assumption on the second stage?
(i.e. Assume that state weakness cannot prevent war)
Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16		0	0	{-1, 0}	{0, 1}	{-1, 0}
3		0	1	{-1, 0}	{~~-1~~, 0}	{0, ~~1~~}
9		1	0	{0, 1}	{~~-1~~, 0}	{~~-1~~, 0}
0		1	1	{0, 1}	{0, 1}	{ 0, 1}

Oil, Civil War, and State Weakness

What if we use the monotone treatment assumption on the second stage?
(i.e. Assume that state weakness cannot prevent war)
Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16		0	0	$\{-1, 0\}$	$\{0, 1\}$	$\{-1, 0\}$
3		0	1	$\{-1, 0\}$	$\{\cancel{-1}, 0\}$	$\{0, \cancel{1}\}$
9		1	0	$\{0, 1\}$	$\{\cancel{-1}, 0\}$	$\{\cancel{-1}, 0\}$
0		1	1	$\{0, 1\}$	$\{0, 1\}$	$\{ 0, 1\}$
						AVG $=$ [-16/28, 2/28]

Oil, Civil War, and State Weakness

What if we use the monotone treatment assumption on the second stage?
(i.e. Assume that state weakness cannot prevent war)
Sample Average Mechanism specific effect on the Treated

# of Obs	OIL	SW	WAR	OIL on SW	SW on WAR	Possible Products
16		0	0	{-1, 0}	{0, 1}	{-1, 0}
3		0	1	{-1, 0}	$\{\cancel{-1}, 0\}$	$\{0, \cancel{1}\}$
9		1	0	{0, 1}	$\{\cancel{-1}, 0\}$	$\{\cancel{-1}, 0\}$
0		1	1	{0, 1}	{0, 1}	{ 0, 1}
						AVG = [-16/28, 2/28]

Notice now that the width is less than 1.