# 5b. Partial Identification

ISS5096 || ECI

Jaewon ("Jay-one") Yoo
National Tsing Hua University

# **1/** Partial Identification

Cf. Manski (2003), *Partial Identification of Probability Distributions*

# No-Assumption Bounds

- **Law of decreasing credibility** (Manski): credibility of inferences decreases with strength of assumptions
    - Idea: pick assumptions and then figure out what you can learn.
    - May not be point identified, but maybe we can **bound** the effect.

# No-Assumption Bounds

- **Law of decreasing credibility** (Manski): credibility of inferences decreases with strength of assumptions
  - Idea: pick assumptions and then figure out what you can learn.
  - May not be point identified, but maybe we can **bound** the effect.
- If $Y$ is bounded $[y_L, y_U]$, $\tau$ logically must be in $[y_L - y_U,\ y_U - y_L]$.

# No-Assumption Bounds

- **Law of decreasing credibility** (Manski): credibility of inferences decreases with strength of assumptions
  - Idea: pick assumptions and then figure out what you can learn.
  - May not be point identified, but maybe we can **bound** the effect.
- If $Y$ is bounded $[y_L, y_U]$, $\tau$ logically must be in $[y_L - y_U,\ y_U - y_L]$.
- Can we improve using data? Rewrite the ATE with $p = \mathbb{P}(D_i = 1)$:

$$\begin{aligned}\tau = {} & \mathbb{E}[Y_i \mid D_i = 1]p + \mathbb{E}[Y_i(1) \mid D_i = 0](1-p) \\ & - \mathbb{E}[Y_i(0) \mid D_i = 1]p - \mathbb{E}[Y_i \mid D_i = 0](1-p)\end{aligned}$$

# No-Assumption Bounds

- **Law of decreasing credibility** (Manski): credibility of inferences decreases with strength of assumptions
  - Idea: pick assumptions and then figure out what you can learn.
  - May not be point identified, but maybe we can **bound** the effect.

- If $Y$ is bounded $[y_L, y_U]$, $\tau$ logically must be in $[y_L - y_U,\ y_U - y_L]$.
- Can we improve using data? Rewrite the ATE with $p = \mathbb{P}(D_i = 1)$:

$$\begin{aligned}\tau = \mathbb{E}[Y_i \mid D_i = 1]p + \mathbb{E}[Y_i(1) \mid D_i = 0](1-p) \\ - \mathbb{E}[Y_i(0) \mid D_i = 1]p - \mathbb{E}[Y_i \mid D_i = 0](1-p)\end{aligned}$$

- Plug in $y_L$ and $y_U$ for the counterfactual means to get bounds for $\tau$:

$$\tau \geq \mathbb{E}[Y_i \mid D_i = 1]p + y_L(1-p) - y_U p - \mathbb{E}[Y_i \mid D_i = 0](1-p)$$
$$\tau \leq \mathbb{E}[Y_i \mid D_i = 1]p + y_U(1-p) - y_L p - \mathbb{E}[Y_i \mid D_i = 0](1-p)$$

  - These bounds have width of $|y_U - y_L|$, i.e., half of the logical bounds.
  - But always will contain 0. Weak assumptions $\rightsquigarrow$ weak inferences.

# Narrowing Bounds with Domain Knowledge

- Domain knowledge can be formalized as **assumptions** that narrow the bounds.

# Narrowing Bounds with Domain Knowledge

- Domain knowledge can be formalized as **assumptions** that narrow the bounds.
- **MTR** (Monotone Treatment Response): treatment doesn't hurt anyone.
  - $Y_i(1) \geq Y_i(0)$ for all $i \rightsquigarrow \tau \geq 0$.
  - E.g., job training won't make you *worse* at finding a job.
  - Domain knowledge pins down the **lower bound**.

# Narrowing Bounds with Domain Knowledge

- Domain knowledge can be formalized as **assumptions** that narrow the bounds.
- **MTR** (Monotone Treatment Response): treatment doesn't hurt anyone.
  - $Y_i(1) \geq Y_i(0)$ for all $i \rightsquigarrow \tau \geq 0$.
  - E.g., job training won't make you *worse* at finding a job.
  - Domain knowledge pins down the **lower bound**.
- **MTS** (Monotone Treatment Selection): positive selection into treatment.
  - $\mathbb{E}[Y(d)|D_i = 1] \geq \mathbb{E}[Y(d)|D_i = 0]$ for $d \in \{0, 1\}$.
  - People with higher baseline outcomes are more likely to seek treatment.
  - Formalizes the **direction** of selection bias as an assumption.

# Narrowing Bounds with Domain Knowledge

- Domain knowledge can be formalized as **assumptions** that narrow the bounds.
- **MTR** (Monotone Treatment Response): treatment doesn't hurt anyone.
  - $Y_i(1) \geq Y_i(0)$ for all $i \rightsquigarrow \tau \geq 0$.
  - E.g., job training won't make you *worse* at finding a job.
  - Domain knowledge pins down the **lower bound**.
- **MTS** (Monotone Treatment Selection): positive selection into treatment.
  - $\mathbb{E}[Y(d)|D_i = 1] \geq \mathbb{E}[Y(d)|D_i = 0]$ for $d \in \{0, 1\}$.
  - People with higher baseline outcomes are more likely to seek treatment.
  - Formalizes the **direction** of selection bias as an assumption.
- Each assumption contributes a *different* piece of information:
  - MTR $\rightsquigarrow$ **lower bound**: the effect is not negative.
  - MTS $\rightsquigarrow$ **upper bound**: the effect is no larger than the observed gap.

# How Does MTS Tighten the Bound?

- Each potential outcome mean has an **observed** and **unobserved** part:

$$\mathbb{E}[Y(1)] = \underbrace{\mathbb{E}[Y_i|D_i = 1]}_{\text{observed}} \cdot p + \underbrace{\mathbb{E}[Y_i(1)|D_i = 0]}_{\text{unobserved}} \cdot (1 - p)$$

$$\mathbb{E}[Y(0)] = \underbrace{\mathbb{E}[Y_i(0)|D_i = 1]}_{\text{unobserved}} \cdot p + \underbrace{\mathbb{E}[Y_i|D_i = 0]}_{\text{observed}} \cdot (1 - p)$$

# How Does MTS Tighten the Bound?

- Each potential outcome mean has an **observed** and **unobserved** part:

$$\mathbb{E}[Y(1)] = \underbrace{\mathbb{E}[Y_i|D_i = 1]}_{\text{observed}} \cdot p + \underbrace{\mathbb{E}[Y_i(1)|D_i = 0]}_{\text{unobserved}} \cdot (1 - p)$$

$$\mathbb{E}[Y(0)] = \underbrace{\mathbb{E}[Y_i(0)|D_i = 1]}_{\text{unobserved}} \cdot p + \underbrace{\mathbb{E}[Y_i|D_i = 0]}_{\text{observed}} \cdot (1 - p)$$

- Upper bound = largest $\tau$ consistent with data and assumptions.
  - Think: what are the largest $\mathbb{E}[Y(1)]$ and smallest $\mathbb{E}[Y(0)]$ possible?
  - This depends on the unobserved counterfactual means.

# How Does MTS Tighten the Bound?

- Each potential outcome mean has an **observed** and **unobserved** part:

$$\mathbb{E}[Y(1)] = \underbrace{\mathbb{E}[Y_i|D_i = 1]}_{\text{observed}} \cdot p + \underbrace{\mathbb{E}[Y_i(1)|D_i = 0]}_{\text{unobserved}} \cdot (1 - p)$$

$$\mathbb{E}[Y(0)] = \underbrace{\mathbb{E}[Y_i(0)|D_i = 1]}_{\text{unobserved}} \cdot p + \underbrace{\mathbb{E}[Y_i|D_i = 0]}_{\text{observed}} \cdot (1 - p)$$

- Upper bound = largest $\tau$ consistent with data and assumptions.
  - Think: what are the largest $\mathbb{E}[Y(1)]$ and smallest $\mathbb{E}[Y(0)]$ possible?
  - This depends on the unobserved counterfactual means.
- MTS replaces **logical bounds** with tighter **data-driven bounds**:

| To find the upper bound... | No assumptions | With MTS |
|---|---|---|
| Largest $\mathbb{E}[Y(1)|D = 0]$ possible | $\leq y_U$ | $\leq \mathbb{E}[Y_i|D_i = 1]$ |
| Smallest $\mathbb{E}[Y(0)|D = 1]$ possible | $\geq y_L$ | $\geq \mathbb{E}[Y_i|D_i = 0]$ |

# How Does MTS Tighten the Bound?

- Each potential outcome mean has an **observed** and **unobserved** part:

$$\mathbb{E}[Y(1)] = \underbrace{\mathbb{E}[Y_i|D_i = 1]}_{\text{observed}} \cdot p + \underbrace{\mathbb{E}[Y_i(1)|D_i = 0]}_{\text{unobserved}} \cdot (1 - p)$$

$$\mathbb{E}[Y(0)] = \underbrace{\mathbb{E}[Y_i(0)|D_i = 1]}_{\text{unobserved}} \cdot p + \underbrace{\mathbb{E}[Y_i|D_i = 0]}_{\text{observed}} \cdot (1 - p)$$

- Upper bound = largest $\tau$ consistent with data and assumptions.
  - Think: what are the largest $\mathbb{E}[Y(1)]$ and smallest $\mathbb{E}[Y(0)]$ possible?
  - This depends on the unobserved counterfactual means.
- MTS replaces **logical bounds** with tighter **data-driven bounds**:

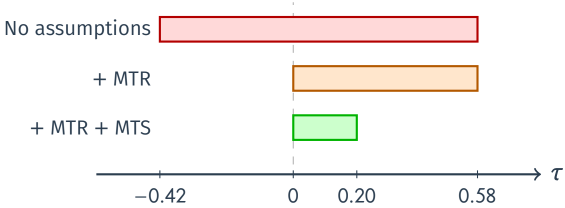
| To find the upper bound... | No assumptions | With MTS |
|---|---|---|
| Largest $\mathbb{E}[Y(1)|D = 0]$ possible | $\leq y_U$ | $\leq \mathbb{E}[Y_i|D_i = 1]$ |
| Smallest $\mathbb{E}[Y(0)|D = 1]$ possible | $\geq y_L$ | $\geq \mathbb{E}[Y_i|D_i = 0]$ |

- For the **lower bound**, directions reverse (smallest $\mathbb{E}[Y(1)]$, largest $\mathbb{E}[Y(0)]$) $\rightsquigarrow$ MTS constraints don't bind; MTR provides the lower bound instead.

# Numerical Example: How Bounds Narrow

Setup: $Y \in [0,1]$, $p = \mathbb{P}(D_i = 1) = 0.4$, $\mathbb{E}[Y_i|D_i = 1] = 0.7$, $\mathbb{E}[Y_i|D_i = 0] = 0.5$.

| Assumptions | Bounds for $\tau$ | Width | What changed? |
|---|---|---|---|
| No assumptions (data only) | $[-0.42,\ 0.58]$ | 1.00 | |
| + MTR | $[\ 0,\ 0.58]$ | 0.58 | lower bound ↑ |
| + MTR + MTS | $[\ 0,\ 0.20]$ | 0.20 | upper bound ↓ |



# Confidence Regions for Bounds

- More general setup:
    - True bounds $[\delta_L, \delta_U]$: the **identification region**.
    - Estimated bounds $[\widehat{\delta}_L, \widehat{\delta}_U]$ with standard errors $\widehat{\text{se}}(\widehat{\delta}_L), \widehat{\text{se}}(\widehat{\delta}_U)$.

# Confidence Regions for Bounds

- More general setup:
  - True bounds $[\delta_L, \delta_U]$: the **identification region**.
  - Estimated bounds $[\widehat{\delta}_L, \widehat{\delta}_U]$ with standard errors $\widehat{\text{se}}(\widehat{\delta}_L), \widehat{\text{se}}(\widehat{\delta}_U)$.

- Goal: find an interval that covers the **true value** $\tau$ with probability $1 - \alpha$:

$$\mathbb{P}(\tau \in [\widehat{\delta}_L, \widehat{\delta}_U]) \geq 1 - \alpha$$

# Confidence Regions for Bounds

- More general setup:
  - True bounds $[\delta_L, \delta_U]$: the **identification region**.
  - Estimated bounds $[\widehat{\delta}_L, \widehat{\delta}_U]$ with standard errors $\widehat{\text{se}}(\widehat{\delta}_L), \widehat{\text{se}}(\widehat{\delta}_U)$.
- Goal: find an interval that covers the **true value** $\tau$ with probability $1 - \alpha$:

$$\mathbb{P}(\tau \in [\widehat{\delta}_L, \widehat{\delta}_U]) \geq 1 - \alpha$$

- Confidence interval:

$$\left[\widehat{\delta}_L - z_{1-\alpha}\,\widehat{\text{se}}(\widehat{\delta}_L), \quad \widehat{\delta}_U + z_{1-\alpha}\,\widehat{\text{se}}(\widehat{\delta}_U)\right]$$

  - If $\tau = \delta_L$ or $\tau = \delta_U$ (boundary): coverage $\longrightarrow 1 - \alpha$.
  - If $\delta_L < \tau < \delta_U$ (interior): coverage $\longrightarrow 1$.
  - Uses one-sided $z_{1-\alpha}$ (not $z_{1-\alpha/2}$) because each bound is one-sided.