**QTM 101 Midterm 2 Cheat Sheet**

**Section 1, 10am-11:15am, Wed, Nov 8th, 2023**

**Name: Jiuru Lyu**

# From Midterm 1

- $\underbrace{\boxed{\text{Estimate}}}_{\text{what we see in data}} = \underbrace{\boxed{\text{Estimand}}}_{\text{what we want to know}} + \text{Bias} + \text{Noise}$
  - $\text{ATE} = \mathbf{E}[Y_1] - \mathbf{E}[Y_0]$

  The only observable effect (Estimate) is $\mathbf{E}[Y_1 \mid T = 1] - \mathbf{E}[Y_0 \mid T = 0]$
- Bias: Systematic error; Not apples-to-apples; Inaccuracy
  $$\text{Bias} = \mathbf{E}[Y_0 \mid T = 1] - \mathbf{E}[Y_0 \mid T = 0]$$
  - Eliminating bias: **randomization**.
- Noise: Random error; $\mathbf{E}[\text{Noise} = 0]$; Imprecision
  - Everything is noisy.
  - Measured by standard error
  - If we repeat our experiment for enough times, our Noise should be $0$.

# Regression

- General formula:
  $$Y_i = \alpha + \beta T_i + \gamma V_i + r_i$$
  - $Y$: dependent variable; $T$: independent variable/treatment; $V$: control variable/confounder
  - $\alpha = \mathbf{E}[Y_0]$: intercept
  - $\beta = \text{ATE}$: slope (estimand)
  - $r$: error term/residual
- Omitted Variable Bias (Confounders):
  - Confounders must affect treatment AND outcome
  - Short (uncontrolled):
    $$Y_i = \alpha + \beta^S T_i + r_i$$
  - Long (controlled):
    $$Y_i = \alpha + \beta^L T_i + \gamma V_i + r_i$$
  - Omitted Variable Bias (OVB):
    $$\text{OVB} = \beta^S - \beta^L.$$
  - If $\beta^S$ and $\beta^L$ are negative, we compare $|\beta^S|$ and $|\beta^L|$. If $|\beta^S| > |\beta^L|$, then we overestimate the negative treatment effect. $\beta^L$ is a **less biased** estimate of the ATE than $\beta^S$.

# Difference-in-Difference

- Key: Parallel Trends. Outcomes must change at the similar rates; lines should have similar slopes.
- Two approaches: Difference over time and different between subjects: Treated minus untreated; post-treatment minus pre-treatment.
- Violation of parallel trends: overestimate and underestimate.

# Regression Discontinuity Design (RDD)

- Running variable, threshold, treatment status
- Local Average Treatment Effect (LATE):

  $\text{LATE} = \mathbf{E}[Y_1 \mid X = \text{threshold}] - \mathbf{E}[Y_0 \mid X = \text{threshold}]$
  - If we change the bin, we will get different LATEs.
  - Different ways to get LATEs:
    - naïve way: choose a bin and compute the bin at both sides of the bin
    - local linear/polynomial: choose a window and run linear/polynomial regressions then compute the limit
- Continuity Assumption: Points must have the same **potential outcomes** at the threshold. **Capacity to manipulate** treatment assignment cannot be related to outcome.

# Noncompliance and Instrument

- Types of Populations:
  - Always Takers
  - Compliers: What we are interested in.
  - Never Takers
  - Defiers: we assume there are no defiers.
- $Z$: our assignment. $T$: whether the subjects receive the treatment.
- $Z \neq T$; $Z$ is randomized $\implies$ $Z = 0$ and $Z = 1$ are apples-to-apples
  - Given $Z = 1$:
    - $T = 1$: always takers + compliers
    - $T = 0$: never takers
  - Given $Z = 0$:
    - $T = 1$: always takers
    - $T = 0$: never takers + compliers
- We can compute the proportion of compliers by
  $$1 - \mathbf{P}(\text{always takers}) - \mathbf{P}(\text{never takers}).$$

- First Stage Treatment:

  $$\begin{aligned}&\mathbf{Avg}[T \mid Z=1] - \mathbf{Avg}[T \mid Z=0]\\ &= \mathbf{P}(\text{always takers}) + \mathbf{P}(\text{compliers}) - \mathbf{P}(\text{always takers})\\ &= \mathbf{P}(\text{compliers})\end{aligned}$$

- Intent-to-Treat (ITT):

  $$\mathbf{Avg}[Y \mid Z=1] - \mathbf{Avg}[Y \mid Z=0]$$

  - ITT under estimates CATE because it averages in always takers and never takers with treatment effect of $0$.
  - We can scale ITT back to the CATE up by the proportion of compliers.

- Complier Average Treatment Effect (CATE):

  $$\text{CATE} = \frac{\text{ITT}}{\mathbf{P}(\text{compliers})}$$

- Assumptions:
  - No defiers
  - Some compliers
  - Exclusion ($Z$ is only associated with $T$)
  - Exogeneity (random or as-if random)

## Useful Phrasings

- *The intention to treat ($Z$) is an admissible instrument for the actual treatment status ($T$) in this randomized experiment with noncompliance.*

  The intention to treat should meet $3$ conditions to be a proper instrumental variable:

  - Relevance + ITT is strongly correlated with actual treatment uptake + evidence: % of compliers
  - Exclusion + evidence (in context): It's hard to argue that the ITT alone causes any sort of other outcome that affects the outcomes than whether or not the individual receives treatment.
  - Exchangeability – because $Z$ is randomized, it cannot correlate with any pre-existing confounder

- *Should we expect that the* CATE *to be a good estimate of the true average treatment effect for* ***all*** *individuals?*
  - CATE is local/only for compliers + context
  - Compliers being different from non-compliers + evidence in context

- *The continuity assumption that is critical to RDD. The relationship between the continuity assumption, "as-if" randomization, and apples-to-apples comparisons.*
  - The continuity assumption: the potential outcomes functions continue smoothly across the threshold – as we transition from what we can see to what we cannot, there are no jumps at the threshold.
  - This assumption bakes in the apples-to-apples comparison at the threshold + explanation with context $Y$ of subjects below the threshold is identical to $Y$ of subjects above the threshold.
  - This is equivalent to random assignment about the threshold, so the LATE is free from bias.

- *Is* LATE *a good estimator of all populations?*
  - Not really.
  - The LATE only applies to individuals at the threshold. Using this design, we learn nothing about the treatment effect for those away from the threshold – it is wholly possible that the potential outcomes for the treated in a world where they are not treated do not follow the "trend" discovered on the left of the threshold.

| | Regression | DiD | RDD | IV |
|---|---|---|---|---|
| When to use | Anytime | Treated and untreated, before and after treatment | Sharp threshold (Treated one side, untreated the other) | Experiments with non-compliance |
| Assumptions | controlling for all confounders (Omitted Variable Bias) | Parallel trends | Continuity (Manipulation of score) | Exogeneity, Exclusion, Compliers, no defiers |
| Name for Estimate | ATE | ATT or DiD | LATE | CATE |

# QTM 110 Midterm 2 Review

November 7, 2023

## 1 Regression

- General formula:
$$Y_i = \alpha + \beta T_i + \gamma V_i + r_i$$
  - $Y$: dependent variable; $T$: independent variable/treatment; $V$: controlled variable (confounder)
  - $\alpha = \mathbf{E}[Y_0]$: intercept
  - $\beta = \text{ATE}$: slope (estimand)
  - $r$: error term/residual
- Omitted Variable Bias (Confounders):
  - Confounders must affect treatment AND outcome
  - Short (uncontrolled):
$$Y_i = \alpha + \beta^S T_i + r_i$$
  - Long (controlled):
$$Y_i = \alpha + \beta^L T_i + \gamma V_i + r_i$$
  - Omitted Variable Bias (OVB):
$$\text{OVB} = \beta^S - \beta^L.$$
  - If $\beta^S > \beta^L$, then we overestimate the treatment effect. $\beta^L$ is a **less biased** estimate of the ATE than $\beta^S$.
  - If $\beta^S$ and $\beta^L$ are negative, we compare $\left|\beta^S\right|$ and $\left|\beta^L\right|$. If $\left|\beta^S\right| > \left|\beta^L\right|$, then we overestimate the negative treatment effect.

## 2 Difference-in-Difference

- Key: Parallel Trends. Outcomes must change at the similar rates; lines should have similar slopes.
- Two approaches: Difference over time and different between subjects.
$$(Y_{\text{treated, post-treatment}} - Y_{\text{treated, pre-treatment}}) - (Y_{\text{untreated, post-treatment}} - Y_{\text{untreated, pre-treatment}})$$
$$(Y_{\text{treated, post-treatment}} - Y_{\text{untreated, post-treatment}}) - (Y_{\text{treated, pre-treatment}} - Y_{\text{untreated, pre-treatment}})$$
- Violation of parallel trends: overestimate and underestimate.

## 3 Regression Discontinuity Design (RDD)

- Running variable
- Local Average Treatment Effect (LATE)
  - If we change the bin, we will get different LATEs.
  - Different ways to get LATEs: naïve way (choose a bin and compute the bin at both sides of the bin), local linear (choose a window and run linear regressions then compute the limit), polynomial (choose window and run polynomial regressions then compute the limit).
- Continuity Assumption: Points must have the same **potential outcomes** at the threshold.
- **Capacity to manipulate** treatment assignment cannot be related to outcome.

# 4 Noncompliance and Instrument

- Types of Populations:
  - Always Takers
  - Compliers: What we are interested in.
  - Never Takers
  - Defiers: we assume there are no defiers.
- $Z$: our assignment. $T$: whether the subjects receive the treatment.
- $Z \neq T$; $Z$ is randomized: $Z = 0$ and $Z = 1$ are apples-to-apples
  - Given $Z = 1$:
    - We could have $T = 1$: always takers + compliers
    - We could have $T = 0$: never takers
  - Given $Z = 0$:
    - We could have $T = 1$: always takers
    - We could have $T = 0$: never takers + compliers
- Since we know
$$\mathbf{P}(\text{compliers}) + \mathbf{P}(\text{always takers}) + \mathbf{P}(\text{never takers}) = 1$$
we can compute the proportion of compliers by
$$1 - \mathbf{P}(\text{always takers}) - \mathbf{P}(\text{never takers}).$$
-
$$\begin{aligned}\mathbf{E}[T \mid Z = 1] &= 1 \times [\mathbf{P}(\text{always takers}) + \mathbf{P}(\text{compliers})] + 0 \times \mathbf{P}(\text{never takers}) \\ &= \mathbf{P}(\text{always takers}) + \mathbf{P}(\text{compliers})\end{aligned}$$
$$\mathbf{E}[T \mid Z = 0] = 1 \times \mathbf{P}(\text{always takers}) + 0 \times [\mathbf{P}(\text{compliers}) + \mathbf{P}(\text{never takers})] = \mathbf{P}(\text{always takers})$$
- First Stage Treatment:
$$\mathbf{Avg}[T \mid Z = 1] - \mathbf{Avg}[T \mid Z = 0] = \mathbf{P}(\text{always takers}) + \mathbf{P}(\text{compliers}) - \mathbf{P}(\text{always takers}) = \mathbf{P}(\text{compliers})$$
- Intent-to-Treat (ITT):
$$\mathbf{Avg}[Y \mid Z = 1] - \mathbf{Avg}[Y \mid Z = 0]$$
  - ITT under estimates CATE because it averages in always takers and never takers with treatment effect of $0$.
  - We can scale ITT back to the CATE up by the proportion of compliers.
- Complier Average Treatment Effect (CATE):
$$\text{CATE} = \frac{\text{ITT}}{\mathbf{P}(\text{compliers})}$$
- Assumptions:
  - No defiers
  - Some compliers
  - Exclusion ($Z$ is only associated with $T$)
  - Exogeneity (random or as-if random)

| | Regression | DiD | RDD | IV |
|---|---|---|---|---|
| When to use | Anytime | Treated and untreated, before and after treatment | Sharp threshold (Treated one side, untreated the other) | Experiments with non-compliance |
| Assumptions | controlling for all confounders (Omitted Variable Bias) | Parallel trends | Continuity (Manipulation of score) | Exogeneity, Exclusion, Compliers, no defiers |
| Name for Estimate | ATE or ATE $\mid X$ | ATT or DiD | LATE | CATE |