# Regression Adjustment and Causality

## Gov 2001: Quantitative Social Science Methods I

Week 13, Lecture 26

Spring 2026

# For Today

## Required Reading

- Angrist & Pischke, §3.3 (pp. 68–91)
- Aronow & Miller, §7.3 (pp. 271–285)

Today: When can regression give us causal effects?

# Roadmap

1. Potential outcomes framework
2. Selection bias
3. Conditional independence assumption (CIA)
4. Regression adjustment under CIA
5. Looking ahead to causal inference

# Part I: Potential Outcomes

# The Potential Outcomes Framework

For each individual $i$:

- $Y_i(1)$ = outcome if treated
- $Y_i(0)$ = outcome if not treated

**Individual treatment effect:**

$$\tau_i = Y_i(1) - Y_i(0)$$

**Fundamental problem of causal inference:**
We only observe one potential outcome per person!
If $D_i = 1$: observe $Y_i(1)$, don't observe $Y_i(0)$
If $D_i = 0$: observe $Y_i(0)$, don't observe $Y_i(1)$

# What We Observe

**Observed outcome:**

$$Y_i = D_i \cdot Y_i(1) + (1 - D_i) \cdot Y_i(0)$$

We see $Y_i(1)$ for treated, $Y_i(0)$ for untreated.
Never both for the same person.

**The counterfactual** $Y_i(0)$ for treated (or $Y_i(1)$ for untreated) is always missing.

# Causal Estimands

**Average Treatment Effect (ATE):**

$$\tau_{ATE} = \mathbb{E}[Y_i(1) - Y_i(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$$

**Average Treatment Effect on the Treated (ATT):**

$$\tau_{ATT} = \mathbb{E}[Y(1) - Y(0)|D = 1]$$

These are **causal** quantities—they involve potential outcomes, not just observed data.

# Part II: Selection Bias

## The Simple Comparison

**Naïve estimator:**

$$\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0]$$

Difference in mean outcomes between treated and untreated.

**What does this equal?**

$$\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0] = \mathbb{E}[Y(1)|D=1] - \mathbb{E}[Y(0)|D=0]$$

We want $\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$, but we get something different.

# Selection Bias

$$
\begin{aligned}
&\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0] \\
&= \mathbb{E}[Y(1)|D=1] - \mathbb{E}[Y(0)|D=0] \\
&= \underbrace{\mathbb{E}[Y(1)|D=1] - \mathbb{E}[Y(0)|D=1]}_{ATT} + \underbrace{\mathbb{E}[Y(0)|D=1] - \mathbb{E}[Y(0)|D=0]}_{\text{Selection Bias}}
\end{aligned}
$$

**Selection bias:** Difference in baseline outcomes between treated and untreated.
If people who select into treatment are systematically different, simple comparisons are biased.

## Example: Campaign Contact and Turnout

**Question:** Effect of GOTV contact on voter turnout?

**Simple comparison:**

$$\text{Turnout (contacted)} - \text{Turnout (not contacted)}$$

**Problem:** Who gets contacted?

- Voters on party lists (already engaged)
- People who answer the door (more social)
- Residents of targeted precincts (already competitive)

These voters would have higher turnout *even without contact.*
$\Rightarrow$ Selection bias (probably positive).

# Part III: Conditional Independence

# Conditional Independence Assumption (CIA)

## Conditional Independence

$$(Y(0), Y(1)) \perp\!\!\!\perp D \mid X$$

**Translation:**
Conditional on observed covariates $X$, treatment is “as good as random.”

**Equivalently:**
Within cells defined by $X$, treated and untreated are comparable.
No selection on unobservables (after conditioning on $X$).

## What CIA Means

**Without conditioning:**

$$\mathbb{E}[Y(0)|D=1] \neq \mathbb{E}[Y(0)|D=0] \quad \text{(selection bias)}$$

**Under CIA:**

$$\mathbb{E}[Y(0)|D=1,X] = \mathbb{E}[Y(0)|D=0,X] = \mathbb{E}[Y(0)|X]$$

Within each $X$-cell, no selection bias!

**Also called:**

- Unconfoundedness
- Selection on observables
- Ignorability (of treatment assignment)

# When Might CIA Hold?

**More plausible when:**

- You observe the key confounders
- Selection into treatment is well understood
- Rich administrative data with many covariates
- Institutional knowledge suggests observables drive selection

**Less plausible when:**

- Important confounders are unobserved
- Selection driven by private information
- Treatment is self-selected based on expected gains

CIA is an assumption—it cannot be tested from data.

# Part IV: Regression Adjustment Under CIA

# Regression Adjustment

**Under CIA:**

$$Y_i = \alpha + \beta D_i + \gamma X_i + \varepsilon_i$$

With CIA holding, the coefficient $\beta$ estimates a causal effect!

**Why?**

CIA $\Rightarrow \mathbb{E}[\varepsilon|D, X] = \mathbb{E}[\varepsilon|X]$
After controlling for $X$, $D$ is uncorrelated with unobservables.
Just like in a randomized experiment within each $X$-cell.

# What Does Regression Estimate Under CIA?

Under CIA + linear CEF:

$$\hat{\beta} \xrightarrow{p} \text{ATE}$$

**But recall:** With heterogeneous effects and varying propensities, regression gives a **variance-weighted** average.

This may differ from simple ATE or ATT.

**For ATE:** May need additional re-weighting (IPW or matching).

# The Overlap Assumption

CIA alone isn't enough. We also need:

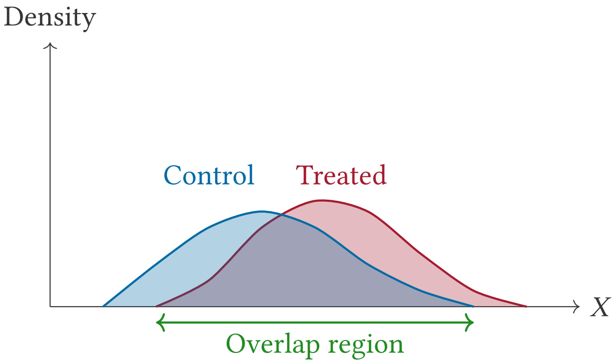
**Overlap (Common Support)**

$$0 < P(D = 1|X) < 1 \quad \text{for all } X$$

**Translation:**
For every value of $X$, there are both treated and untreated units.

**Why?**
Can't compare treated and untreated if one group is empty in some $X$-cells.

## Overlap Illustrated




Regression adjustment works in the overlap region.

# Part V: Looking Ahead

# Limitations of Regression Adjustment

**1. CIA is strong and untestable**
We can't know if we've controlled for everything relevant.

**2. Functional form matters**
Linear regression may not capture complex relationships.

**3. Variance weighting**
May not give the estimand you want.

**4. Extrapolation beyond overlap**
Regression can give answers where data is thin.

# Beyond Regression: Other Causal Methods

When CIA is plausible, regression is one option. Others include:

**Matching:** Find similar units in treatment and control groups.
**Inverse Probability Weighting (IPW):** Re-weight to balance covariates.
**Doubly Robust:** Combine regression and IPW.

When CIA fails, need other strategies:
**Instrumental Variables, Regression Discontinuity, Difference-in-Differences, Synthetic Control...**

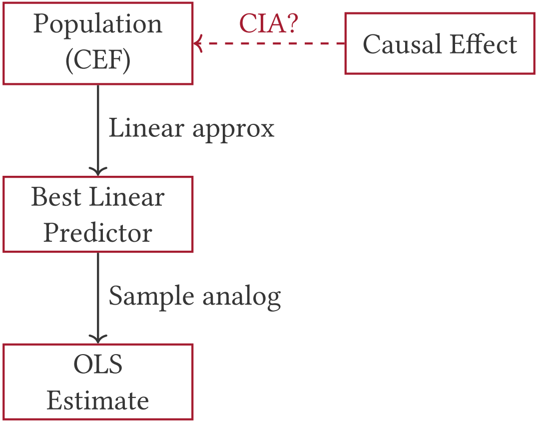
(Topics for Gov 2002 and causal inference courses!)

# Course Summary

**What we've learned:**

1. **Probability**: The language for describing uncertainty
2. **Random variables**: Expected value, variance, distributions
3. **The CEF**: $\mathbb{E}[Y|X]$ is our target
4. **Sampling and CLT**: How samples inform about populations
5. **Estimation and inference**: Point estimates, SEs, CIs, tests
6. **BLP and OLS**: Linear approximation to CEF
7. **Multiple regression**: Controlling for covariates, FWL, OVB
8. **Inference**: Robust and clustered standard errors
9. **Causality**: When regression gives causal effects

## The Big Picture




OLS estimates the BLP, which approximates the CEF.
Under CIA, the CEF difference equals the causal effect.

# Final Thoughts

**Regression is a tool**, not a magic wand.

- It always estimates *something* (the BLP)
- Whether that's *causal* depends on assumptions
- Understanding what you're estimating is crucial

**The goal of this course:**
Understand the machinery of statistical inference well enough to use it critically and responsibly.

# Summary

**Today's key points:**

1. Potential outcomes define causal effects
2. Selection bias arises when treated/control differ at baseline
3. CIA: conditional on $X$, treatment is as-if random
4. Under CIA, regression adjustment can identify causal effects
5. But CIA is untestable—domain knowledge required

Regression can give causal estimates
under conditional independence (CIA).

But CIA requires controlling for all confounders—
an assumption that cannot be tested.

Critical thinking about assumptions
is what separates good research from bad.

## Thank you for a great semester!

Good luck on the final.

Keep learning, keep questioning, keep thinking.