Structural Equation Modeling

Kosuke Imai

Princeton University

POL572 Quantitative Analysis II
Spring 2016

Causal Mediation Analysis

Quantitative Research and Causal Mechanisms

- Causal inference is a central goal of scientific research
- Scientists care about *causal mechanisms*, not just *causal effects*

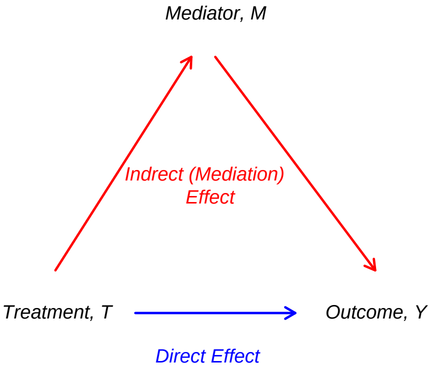
- Randomized experiments often only determine whether the treatment causes changes in the outcome
- Not how and why the treatment affects the outcome
- Common criticism of experiments and statistics:

black box view of causality

- Qualitative research uses process tracing
- **Question:** How can quantitative research be used to identify causal mechanisms?

Direct and Indirect Effects

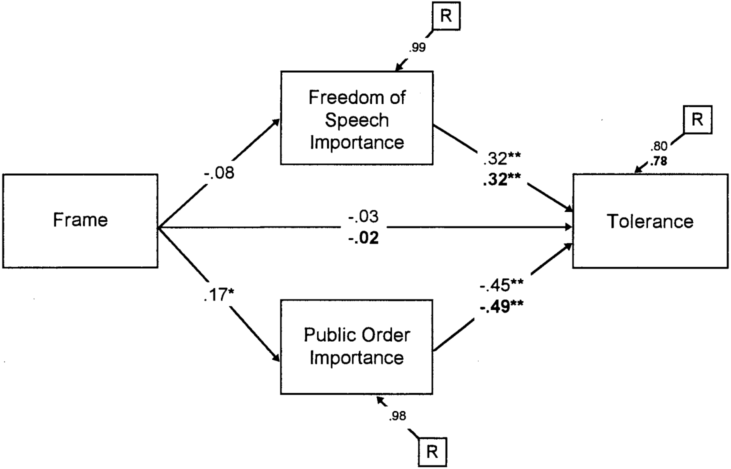
- Causal mediation analysis:



- Direct and indirect effects; intermediate and intervening variables
- READING: Imai, Keele, Tingley, and Yamamoto (2012). *APSR*

Causal Mediation Analysis in American Politics

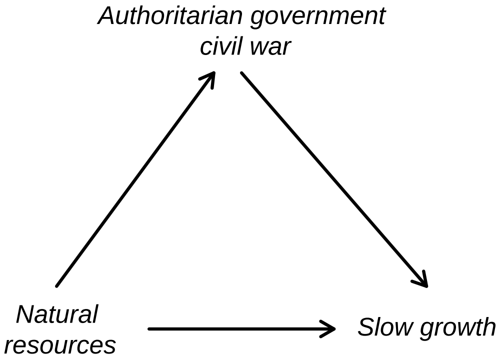
- Media framing experiment in Nelson *et al.* (*APSR*, 1998)



- Path analysis, structural equation modeling

Causal Mediation Analysis in Comparative Politics

- Resource curse thesis



- Causes of civil war: Fearon and Laitin (*APSR*, 2003)

Causal Mediation Analysis in International Relations

- The literature on international regimes and institutions
- Krasner (*International Organization*, 1982)

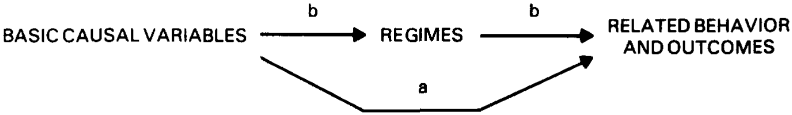


Figure 2

- Power and interests are mediated by regimes

Potential Outcomes Framework for Mediation

- Binary treatment: T_i
- Pre-treatment covariates: X_i

- Potential mediators: $M_i(t)$
- Observed mediator: $M_i = M_i(T_i)$

- Potential outcomes: $Y_i(t, m)$
- Observed outcome: $Y_i = Y_i(T_i, M_i(T_i))$

- Again, **only one potential outcome can be observed per unit**

Causal Mediation Effects

- Total causal effect:

$$\tau_i \equiv Y_i(1, M_i(1)) - Y_i(0, M_i(0))$$

- Causal mediation (Indirect) effects:

$$\delta_i(t) \equiv Y_i(t, M_i(1)) - Y_i(t, M_i(0))$$

- Causal effect of the treatment-induced change in M_i on Y_i
- Change the mediator from $M_i(0)$ to $M_i(1)$ while holding the treatment constant at t
- Represents the mechanism through M_i

Total Effect = Indirect Effect + Direct Effect

- Direct effects:

$$\zeta_i(t) \equiv Y_i(1, M_i(t)) - Y_i(0, M_i(t))$$

- Causal effect of T_i on Y_i, holding mediator constant at its potential value that would be realized when $T_i = t$
- Change the treatment from 0 to 1 while holding the mediator constant at $M_i(t)$
- Represents all mechanisms other than through M_i

- Total effect = mediation (indirect) effect + direct effect:

$$\tau_i = \delta_i(t) + \zeta_i(1 - t)$$

What Does the Observed Data Tell Us?

- Quantity of Interest: Average causal mediation effects (ACME)

$$\bar{\delta}(t) \equiv \mathbb{E}(\delta_i(t)) = \mathbb{E}\{Y_i(t, M_i(1)) - Y_i(t, M_i(0))\}$$

- Average direct effects ($\bar{\zeta}(t)$) are defined similarly

- $Y_i(t, M_i(t))$ is observed but $Y_i(t, M_i(t'))$ can never be observed
- We have an identification problem

$\Longrightarrow$ Need additional assumptions to make progress

Identification under Sequential Ignorability

- Proposed identification assumption: Sequential Ignorability (SI)

$$\{Y_i(t', m), M_i(t)\} \perp\!\!\!\perp T_i \mid X_i = x, \qquad (1)$$

$$Y_i(t', m) \perp\!\!\!\perp M_i(t) \mid T_i = t, X_i = x \qquad (2)$$

- (1) is guaranteed to hold in a standard experiment
- (2) does **not** hold unless X_i includes all confounders
- Like any ignorability assumption, the assumption is very strong
    1. No unmeasured pre-treatment confounder
    2. No measured and unmeasured post-treatment confounder

Under SI, ACME is nonparametrically identified:

$$\int\int \mathbb{E}(Y_i \mid M_i, T_i = t, X_i) \{dP(M_i \mid T_i = 1, X_i) - dP(M_i \mid T_i = 0, X_i)\} \, dP(X_i)$$

The Linear Structural Equation Model

$$
\begin{aligned}
Y_i &= \alpha_1 + \beta_1 T_i + X_i^\top \xi_1 + \epsilon_{1i}, && (1)\\
M_i &= \alpha_2 + \beta_2 T_i + X_i^\top \xi_2 + \epsilon_{2i}, && (2)\\
Y_i &= \alpha_3 + \beta_3 T_i + \gamma M_i + X_i^\top \xi_3 + \epsilon_{3i}. && (3)
\end{aligned}
$$

- Equation (1) is redundant given equations (2) and (3):
  $\alpha_1 = \alpha_3 + \alpha_2\gamma$, $\beta_1 = \beta_2\gamma + \beta_3$, $\xi_1 = \xi_2\gamma + \xi_3$, $\epsilon_{1i} = \gamma\epsilon_{2i} + \epsilon_{3i}$

- Potential outcomes notation:

$$
\begin{aligned}
M_i(T_i) &= \alpha_2 + \beta_2 T_i + X_i^\top \xi_2 + \epsilon_{2i},\\
Y_i(T_i, M_i(T_i)) &= \alpha_3 + \beta_3 T_i + \gamma M_i(T_i) + X_i^\top \xi_3 + \epsilon_{3i}.
\end{aligned}
$$

- Under SI and linearity:
  1. $\mathbb{E}(\epsilon_{2i} \mid T_i, X_i) = \mathbb{E}(\epsilon_{2i} \mid X_i) = f_2(X_i) = X_i^\top \xi_2^*$
  2. $\mathbb{E}(\epsilon_{3i} \mid T_i, M_i, X_i) = \mathbb{E}(\epsilon_{3i} \mid X_i) = f_3(X_i) = X_i^\top \xi_3^*$

Estimation of Average Causal Mediation Effects

- Redefine:
  1. $\epsilon_{2i} = X_i^\top \xi_2^* + \epsilon_{2i}^*$ where $\mathbb{E}(\epsilon_{2i}^* \mid X_i) = 0$
  2. $\epsilon_{3i} = X_i^\top \xi_3^* + \epsilon_{3i}^*$ where $\mathbb{E}(\epsilon_{3i}^* \mid X_i) = 0$
- Then, SI implies $\mathrm{Cor}(\epsilon_{2i}^*, \epsilon_{3i}^*) = 0$
- Non-zero correlation $\Longrightarrow$ M_i is correlated with ϵ_{3i} given T_i and X_i

- Under the LSEM and SI,
  1. $\bar{\delta}(1) = \bar{\delta}(0) = \beta_2\gamma$
  2. $\bar{\zeta}(1) = \bar{\zeta}(0) = \beta_3$
  3. $\tau = \beta_2\gamma + \beta_3$
- Exact variance: $\mathbb{V}(\hat{\beta}_2\hat{\gamma}) = \beta_2^2\mathbb{V}(\hat{\gamma}) + \gamma^2\mathbb{V}(\hat{\beta}_2) + \mathbb{V}(\hat{\gamma})\mathbb{V}(\hat{\beta}_2)$
- Asymptotic variance: $\mathbb{V}(\hat{\beta}_2\hat{\gamma}) \approx \beta_2^2\mathbb{V}(\hat{\gamma}) + \gamma^2\mathbb{V}(\hat{\beta}_2)$
- No interaction assumption: $\bar{\delta}(1) = \bar{\delta}(0)$
- Can be relaxed via an interaction term in equation (3): T_iM_i

The Delta Method

- Taylor's Theorem: If $f(x)$ is $n+1$ times differentiable and $f^{(n)}$ is continuous, then for each x there exists $c \in (x_0, x)$ such that,

$$f(x) = f(x_0) + \underbrace{\sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k}_{Taylor\ series} + \underbrace{\frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{(n+1)}}_{Lagrange\ remainder}$$

- Univariate: Suppose that $\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$. Then, $\sqrt{n}(f(X_n) - f(\theta)) \xrightarrow{d} \mathcal{N}(0, \sigma^2[f^{(1)}(\theta)]^2)$

- Multivariate: Suppose that $\sqrt{n}(X_n - \theta) \xrightarrow{d} \mathcal{N}(0, \Sigma)$. Then, $\sqrt{n}(f(X_n) - f(\theta)) \xrightarrow{d} \mathcal{N}(0, [f^{(1)}(\theta)]^\top \Sigma f^{(1)}(\theta))$

General Estimation Algorithm

Based on the nonparametric identification result under SI:

1. Model outcome and mediator
    - Outcome model: $p(Y_i \mid T_i, M_i, X_i)$
    - Mediator model: $p(M_i \mid T_i, X_i)$
    - These models can be of any form (linear or nonlinear, semi- or nonparametric, with or without interactions)
2. Predict mediator for both treatment values ($M_i(1)$, $M_i(0)$)
3. Predict outcome by first setting $T_i = 1$ and $M_i = M_i(0)$, and then $T_i = 1$ and $M_i = M_i(1)$
4. Compute the average difference between two outcomes to obtain a consistent estimate of ACME
5. Monte Carlo simulation or bootstrap to estimate uncertainty (more on these methods later)

Need for Sensitivity Analysis

- The sequential ignorability assumption is often too strong
- Need to assess the robustness of findings via sensitivity analysis
- Question: How large a departure from the key assumption must occur for the conclusions to no longer hold?

- Sensitivity analysis by assuming

$$\{Y_i(t', m), M_i(t)\} \perp\!\!\!\perp T_i \mid X_i = x$$

but not

$$Y_i(t', m) \perp\!\!\!\perp M_i \mid T_i = t, X_i = x$$

- Possible existence of unobserved *pre-treatment* confounder

Sensitivity Analysis

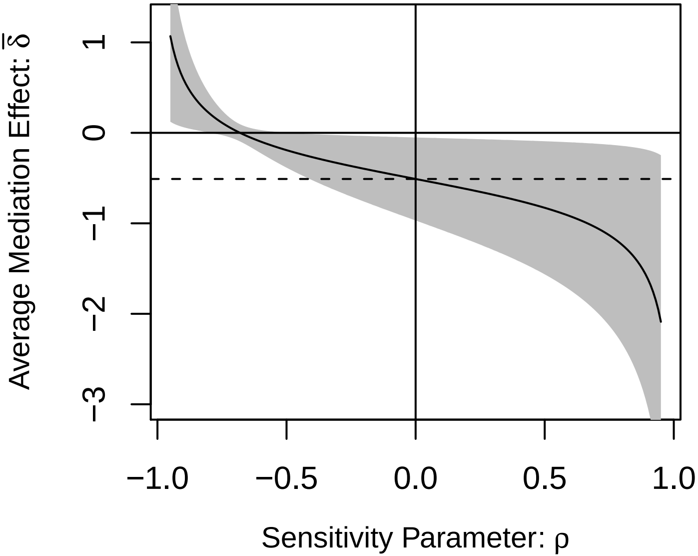
- Sensitivity parameter: $\rho \equiv \mathrm{Corr}(\epsilon_{2i}, \epsilon_{3i})$
- Sequential ignorability implies $\rho = 0$
- Set ρ to different values and see how mediation effects change
- Identification with a given error correlation

$$\bar{\delta}(0) = \bar{\delta}(1) = \beta_2 \left(\frac{\sigma_{12}}{\sigma_2^2} - \frac{\rho}{\sigma_2} \sqrt{\frac{1}{1-\rho^2}\left(\sigma_1^2 - \frac{\sigma_{12}^2}{\sigma_2^2}\right)} \right),$$

where $\sigma_j^2 \equiv \mathbb{V}(\epsilon_{ij})$ for $j = 1, 2$ and $\sigma_{12} \equiv \mathrm{Cov}(\epsilon_{1i}, \epsilon_{2i})$.

- When do my results go away completely?
- $\bar{\delta}(t) = 0$ if and only if $\rho = \mathrm{Corr}(\epsilon_{1i}, \epsilon_{2i})$ (easy to compute!)
- Interpretation of ρ can be based on R^2 statistics

An Example: Nelson *et al.* (1998) *APSR*



Beyond Sequential Ignorability

- Sensitivity analysis may be unsatisfactory
- Existence of post-treatment confounders: multiple mediators

- What if we get rid of the assumption altogether?
- Under a standard design, even the sign of ACME is unidentified
- Can we do any better?

- Develop alternative designs for more credible and powerful inference
- Designs feasible when the mediator can be directly or indirectly manipulated
- Statistical methods for handling multiple mediators

Instrumental Variables

Instrumental Variables

- The Model: $Y = \mathbf{X}\beta + \epsilon$ with $\mathbb{E}(\epsilon_i) = 0$ and $\mathbb{V}(\epsilon_i) = \sigma^2$
- Endogeneity: $\mathbf{X} \not\perp\!\!\!\perp \epsilon$ where $\mathbf{X}$ is $n \times K$
- Instruments: $\mathbf{Z} \perp\!\!\!\perp \epsilon$ where $\mathbf{Z}$ is $n \times L$
- Rank condition: $\mathbf{Z}^\top \mathbf{X}$ and $\mathbf{Z}^\top \mathbf{Z}$ have full rank
- Identification
  1. $K = L$: just-identified
  2. $K < L$: over-identified
  3. $K > L$: under-identified

- The IV estimator:

$$\hat{\beta}_{\text{IV}} \equiv (\mathbf{X}^\top \mathbf{Z}(\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Z}(\mathbf{Z}^\top \mathbf{Z})^{-1} \mathbf{Z}^\top Y$$

Geometry of Instrumental Variables

- Projection matrix (onto $\mathcal{S}(\mathbf{Z})$): $\mathbf{P_Z} = \mathbf{Z}(\mathbf{Z}^\top \mathbf{Z})^{-1}\mathbf{Z}^\top$
- "Purge" endogeneity: $\widehat{\mathbf{X}} = \mathbf{P_Z X}$
- Since $\mathbf{P_Z} = \mathbf{P_Z^\top}$ and $\mathbf{P_Z P_Z} = \mathbf{P_Z}$, we have

$$\hat{\beta}_{\mathrm{IV}} = (\widehat{\mathbf{X}}^\top \widehat{\mathbf{X}})^{-1} \widehat{\mathbf{X}}^\top Y$$

- Two stage least squares:
  1. Regress $\mathbf{X}$ on $\mathbf{Z}$ and obtain the fitted values $\widehat{\mathbf{X}}$
  2. Regress Y on $\widehat{\mathbf{X}}$
- We do not assume the linearity of $\mathbf{X}$ in $\mathbf{Z}$

Asymptotic Inference

- A familiar (by now!) trick:

$$
\begin{aligned}
\hat{\beta}_{\text{IV}} - \beta &= (\mathbf{X}^\top \mathbf{Z}(\mathbf{Z}^\top \mathbf{Z})^{-1}\mathbf{Z}^\top \mathbf{X})^{-1}\mathbf{X}^\top \mathbf{Z}(\mathbf{Z}^\top \mathbf{Z})^{-1}\mathbf{Z}^\top \epsilon \\
&= \left\{ \left(\frac{1}{n}\sum_{i=1}^n X_i Z_i^\top\right) \left(\frac{1}{n}\sum_{i=1}^n Z_i Z_i^\top\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^n Z_i X_i^\top\right) \right\}^{-1} \\
&\quad \times \left(\frac{1}{n}\sum_{i=1}^n X_i Z_i^\top\right) \left(\frac{1}{n}\sum_{i=1}^n Z_i Z_i^\top\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^n Z_i \epsilon_i\right)
\end{aligned}
$$

- Thus, $\hat{\beta}_{\text{IV}} \xrightarrow{p} \beta$
- Under the homoskedasticity, $\mathbb{V}(\epsilon \mid \mathbf{Z}) = \sigma^2 \mathbf{I}_n$:

$$
\sqrt{n}(\hat{\beta}_{\text{IV}} - \beta) \xrightarrow{d} \mathcal{N}(0,\ \sigma^2[\mathbb{E}(X_i Z_i^\top)\{\mathbb{E}(Z_i Z_i^\top)\}^{-1}\mathbb{E}(Z_i X_i^\top)]^{-1})
$$

Residuals and Robust Standard Error

- $\hat{\epsilon} = Y - \mathbf{X}\hat{\beta}_{\text{IV}}$ and not $\hat{\epsilon} \neq Y - \widehat{\mathbf{X}}\hat{\beta}_{\text{IV}}$
- Under homoskedasticity: $\hat{\sigma}^2 \equiv \frac{\|\hat{\epsilon}\|^2}{n-K} \xrightarrow{p} \sigma^2$
- $\widehat{\mathbb{V}(\hat{\beta}_{\text{IV}})} = \hat{\sigma}^2\{\mathbf{X}^\top\mathbf{Z}(\mathbf{Z}^\top\mathbf{Z})^{-1}\mathbf{Z}^\top\mathbf{X}\}^{-1} = \hat{\sigma}^2(\widehat{\mathbf{X}}^\top\widehat{\mathbf{X}})^{-1}$
- Sandwich heteroskedasticity consistent estimator:

$$\begin{aligned}
\text{bread} &= \{\mathbf{X}^\top\mathbf{Z}(\mathbf{Z}^\top\mathbf{Z}_i)^{-1}\mathbf{Z}^\top\mathbf{X}\}^{-1}\mathbf{X}^\top\mathbf{Z}(\mathbf{Z}^\top\mathbf{Z})^{-1} \\
\text{meat} &= \mathbf{Z}^\top \text{diag}(\hat{\epsilon}_i^2)\mathbf{Z} \left(= \sum_{i=1}^{n} \hat{\epsilon}_i^2 Z_i Z_i^\top \right) \\
\text{bread meat bread}^\top &= (\widehat{\mathbf{X}}^\top\widehat{\mathbf{X}})^{-1}\widehat{\mathbf{X}}^\top \text{diag}(\hat{\epsilon}_i^2)\widehat{\mathbf{X}}(\widehat{\mathbf{X}}^\top\widehat{\mathbf{X}})^{-1}
\end{aligned}$$

- Robust standard error for clustering and auto-correlation

Relationship with Linear Structural Equation Modeling

- The model:

$$\begin{aligned} T_i &= \alpha_2 + \beta_2 Z_i + X_i^\top \xi_2 + \epsilon_{i2}, \\ Y_i &= \alpha_3 + \underbrace{\beta_3}_{=0} Z_i + \gamma T_i + X_i^\top \xi_3 + \epsilon_{i3}. \end{aligned}$$

- Exclusion restriction (no direct effect): $\beta_3 = 0$
- $\epsilon_{i2} \not\perp\!\!\!\perp \epsilon_{i3} \mid X_i = x$ for some x
- Ignorability is assumed for Z_i
- Sequential ignorability is not assumed for T_i
- What is the causal interpretation of the IV method?

Partial Compliance in Randomized Experiments

- Unable to force all experimental subjects to take the (randomly) assigned treatment/control
- Intention-to-Treat (ITT) effect $\neq$ treatment effect
- Selection bias: self-selection into the treatment/control groups

- Political information bias: effects of campaign on voting behavior
- Ability bias: effects of education on wages
- Healthy-user bias: effects of exercises on blood pressure

- Encouragement design: randomize the encouragement to receive the treatment rather than the receipt of the treatment itself

Potential Outcomes Notation

- Randomized encouragement: $Z_i \in \{0, 1\}$
- Potential treatment variables: $(T_i(1), T_i(0))$
  1. $T_i(z) = 1$: would receive the treatment if $Z_i = z$
  2. $T_i(z) = 0$: would not receive the treatment if $Z_i = z$
- Observed treatment receipt indicator: $T_i = T_i(Z_i)$
- Observed and potential outcomes: $Y_i = Y_i(Z_i, T_i(Z_i))$
- Can be written as $Y_i = Y_i(Z_i)$

- No interference assumption for $T_i(Z_i)$ and $Y_i(Z_i, T_i)$
- Randomization of encouragement:

$$(Y_i(1), Y_i(0), T_i(1), T_i(0)) \perp\!\!\!\perp Z_i$$

- But $(Y_i(1), Y_i(0)) \not\perp\!\!\!\perp T_i \mid Z_i = z$

Principal Stratification Framework

- Four principal strata (latent types):
  - compliers $(T_i(1), T_i(0)) = (1, 0)$,
  - non-compliers $\left\{ \begin{array}{cc} always-takers & (T_i(1), T_i(0)) = (1, 1), \\ never-takers & (T_i(1), T_i(0)) = (0, 0), \\ defiers & (T_i(1), T_i(0)) = (0, 1) \end{array} \right.$

- Observed and principal strata:

	$Z_i = 1$	$Z_i = 0$
$T_i = 1$	Complier/Always-taker	Defier/Always-taker
$T_i = 0$	Defier/Never-taker	Complier/Never-taker

Instrumental Variables

- Randomized encouragement as an instrument for the treatment
- Two additional assumptions
  1. Monotonicity: No defiers

$$T_i(1) \geq T_i(0) \quad \text{for all } i.$$

  2. Exclusion restriction: Instrument (encouragement) affects outcome only through treatment

$$Y_i(1, t) \; = \; Y_i(0, t) \quad \text{for } t = 0, 1$$

  Zero ITT effect for always-takers and never-takers

- ITT effect decomposition:

$$\begin{aligned} \text{ITT} &= \text{ITT}_c \times \Pr(\text{compliers}) + \text{ITT}_a \times \Pr(\text{always} - \text{takers}) \\ &\quad + \text{ITT}_n \times \Pr(\text{never} - \text{takers}) \\ &= \text{ITT}_c \times \Pr(\text{compliers}) \end{aligned}$$

IV Estimand and Interpretation

- IV estimand:

$$\begin{aligned}
\text{ITT}_c &= \frac{\text{ITT}}{\Pr(\text{compliers})} \\
&= \frac{\mathbb{E}(Y_i \mid Z_i = 1) - \mathbb{E}(Y_i \mid Z_i = 0)}{\mathbb{E}(T_i \mid Z_i = 1) - \mathbb{E}(T_i \mid Z_i = 0)} \\
&= \frac{\text{Cov}(Y_i, Z_i)}{\text{Cov}(T_i, Z_i)}
\end{aligned}$$

- ITT_c = Complier Average Treatment Effect (CATE)
- CATE $\neq$ ATE unless ATE for noncompliers equals CATE
- Different encouragement (instrument) yields different compliers

Asymptotic Inference

- Wald estimator: $\widehat{\text{IV}}_W \equiv \frac{\widehat{\text{Cov}}(Y_i, Z_i)}{\widehat{\text{Cov}}(T_i, Z_i)} = \frac{\widehat{\text{ITT}}_Y}{\widehat{\text{ITT}}_T}$
- Consistency: $\widehat{\text{IV}}_W \xrightarrow{p} \text{CATE} = \text{ITT}_c$
- Asymptotic variance:

$$\begin{aligned} \mathbb{V}(\widehat{\text{IV}}_W) \approx & \frac{1}{\text{ITT}_T^4}\Big\{\text{ITT}_T^2\, \mathbb{V}(\widehat{\text{ITT}}_Y) + \text{ITT}_Y^2\, \mathbb{V}(\widehat{\text{ITT}}_T) \\ & -2\, \text{ITT}_Y\, \text{ITT}_T\, \text{Cov}(\widehat{\text{ITT}}_Y, \widehat{\text{ITT}}_T)\Big\}. \end{aligned}$$

Violations of IV Assumptions

- Violation of exclusion restriction:

$$\text{Large sample bias} \quad = \quad \text{ITT}_{\text{noncomplier}} \times \frac{\Pr(\text{noncomplier})}{\Pr(\text{complier})}$$

- Weak encouragement (instruments)
- Direct effects of encouragement; failure of randomization, alternative causal paths

- Violation of monotonicity:

$$\text{Large sample bias} \quad = \quad \frac{\{\text{CATE} + \text{ITT}_{\text{defier}}\} \Pr(\text{defier})}{\Pr(\text{complier}) - \Pr(\text{defier})}$$

- Proportion of defiers
- Heterogeneity of causal effects

An Example: Testing Habitual Voting

- Randomized encouragement to vote in an election
- Treatment: turnout in the election
- Outcome: turnout in the next election

- Monotonicity: Being contacted by a canvasser would *never* discourage anyone from voting
- Exclusion restriction: being contacted by a canvasser in this election has no effect on turnout in the next election other than through turnout in this election
- CATE: Habitual voting for those who would vote if and only if they are contacted by a canvasser in this election

Multi-valued Treatment

- Two stage least squares regression:

$$\begin{aligned} T_i &= \alpha_2 + \beta_2 Z_i + \eta_i, \\ Y_i &= \alpha_3 + \gamma T_i + \epsilon_i. \end{aligned}$$

- Binary encouragement and binary treatment,
  - $\hat{\gamma} = \widehat{\text{CATE}}$ (no covariate)
  - $\hat{\gamma} \xrightarrow{p} \text{CATE}$ (with covariates)
- Binary encouragement multi-valued treatment
- Monotonicity: $T_i(1) \geq T_i(0)$
- Exclusion restriction: $Y_i(1, t) = Y_i(0, t)$ for each $t = 0, 1, \ldots, K$

- Estimator

$$\hat{\gamma}_{TSLS} \xrightarrow{p} \frac{\mathrm{Cov}(Y_i, Z_i)}{\mathrm{Cov}(T_i, Z_i)} = \frac{\mathbb{E}(Y_i(1) - Y_i(0))}{\mathbb{E}(T_i(1) - T_i(0))}$$

$$= \sum_{k=0}^{K} \sum_{j=k+1}^{K} w_{jk} \mathbb{E}\left(\frac{Y_i(1) - Y_i(0)}{j - k} \;\middle|\; T_i(1) = j, T_i(0) = k\right)$$

where w_{jk} is the weight, which sums up to one, defined as,

$$w_{jk} = \frac{(j - k)\Pr(T_i(1) = j, T_i(0) = k)}{\sum_{k'=0}^{K} \sum_{j'=k'+1}^{K} (j' - k')\Pr(T_i(1) = j', T_i(0) = k')}.$$

- Easy interpretation under the constant additive effect assumption for every complier type
- Assume encouragement induces at most only one additional dose
- Then, $w_k = \Pr(T_i(1) = k, T_i(0) = k - 1)$

Fuzzy Regression Discontinuity Design

- Sharp regression discontinuity design: $T_i = \mathbf{1}\{X_i \geq c\}$
- What happens if we have noncompliance?
- Forcing variable as an instrument: $Z_i = \mathbf{1}\{X_i \geq c\}$
- Potential outcomes: $T_i(z)$ and $Y_i(z, t)$

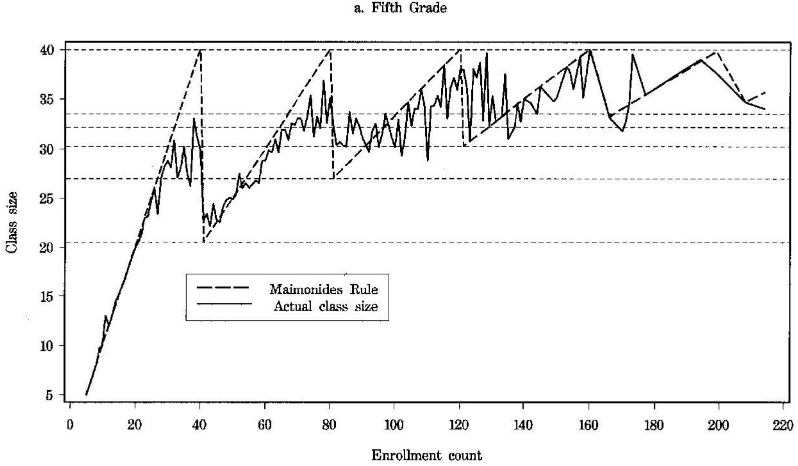
- Monotonicity: $T_i(1) \geq T_i(0)$
- Exclusion restriction: $Y_i(0, t) = Y_i(1, t)$
- $\mathbb{E}(T_i(z) \mid X_i = x)$ and $\mathbb{E}(Y_i(z, T_i(z)) \mid X_i = x)$ are continuous in x
- Estimand: $\mathbb{E}(Y_i(1, T_i(1)) - Y_i(0, T_i(0)) \mid \textit{Complier}, X_i = c)$
- Estimator:

$$\frac{\lim_{x \downarrow c} \mathbb{E}(Y_i \mid X_i = x) - \lim_{x \uparrow c} \mathbb{E}(Y_i \mid X_i = x)}{\lim_{x \downarrow c} \mathbb{E}(T_i \mid X_i = x) - \lim_{x \uparrow c} \mathbb{E}(T_i \mid X_i = x)}$$

- Disadvantage: external validity

An Example: Class Size Effect (Angrist and Lavy)

- Effect of class-size on student test scores
- Maimonides' Rule: Maximum class size $= 40$



Concluding Remarks

- Causal mediation analysis: exploring causal mechanisms
- Even in randomized experiments, an additional (untestable) assumption is required
- Importance of sensitivity analysis

- Instrumental variables in randomized experiments: dealing with partial compliance
- Additional (untestable) assumptions are required
- ITT vs. CATE

- Instrumental variables in observational studies: dealing with selection bias
- Validity of instrumental variables requires rigorous justification
- Tradeoff between internal and external validity