Statistical Analysis of Causal Mechanisms

Kosuke Imai

Princeton University

April 13, 2009

Papers and Software

- Collaborators: Luke Keele, Dustin Tingley, and Teppei Yamamoto
- "Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects." available at http://imai.princeton.edu/research/mediation.html
- "A General Approach to Causal Mediation Analysis"
- "Causal Mediation Analysis in R"
- R package `mediation`

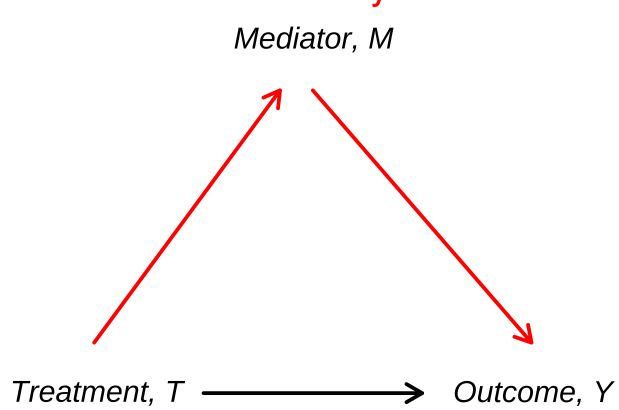
Statistics and Causal Mechanisms

- Causal inference is a central goal of social science and public policy research
- Randomized experiments are seen as **gold standard**
- Design and analyze observational studies to *replicate* experiments
- But, experiments are a **black box**
- Can only tell *whether* the treatment causally affects the outcome
- Not *how* and *why* the treatment affects the outcome
- Qualitative research uses process tracing

- How can quantitative research be used to identify causal mechanisms?

Overview of the Talk

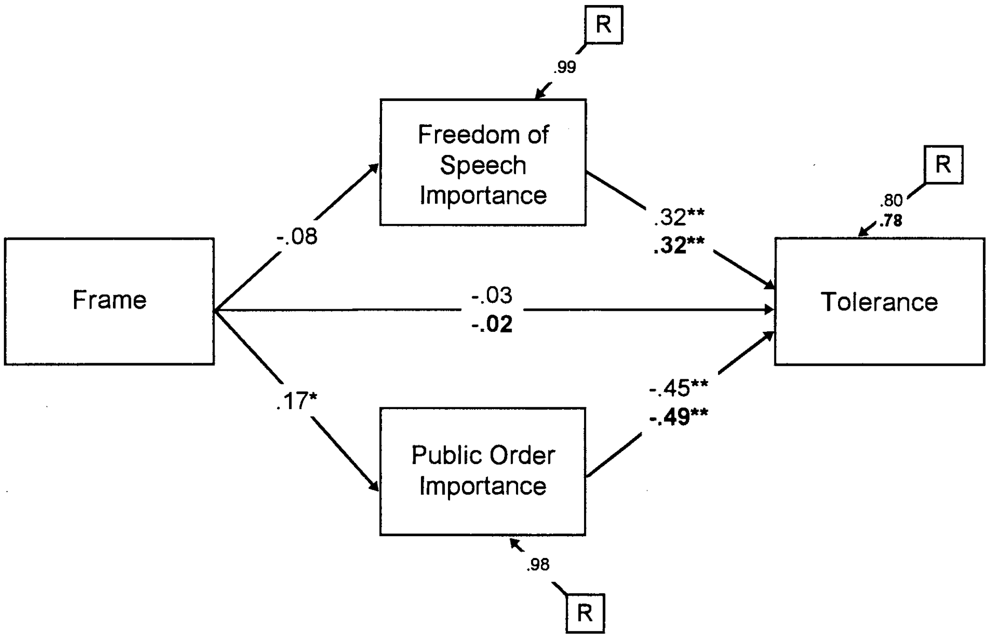
- **Goal:** Convince you that statistics *can* play a role in identifying causal mechanisms
- **Method:** Causal Mediation Analysis



- Direct and indirect effects; intermediate and intervening variables
- Path analysis, structural equation modeling

Causal Mediation Analysis in American Politics

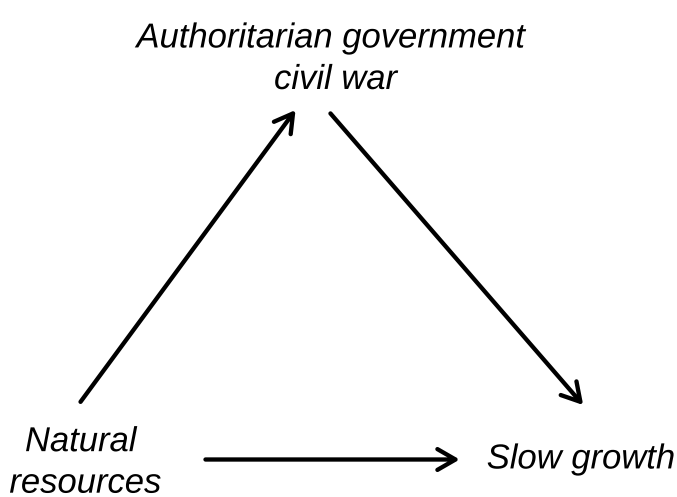
- The political psychology literature on media framing
- Nelson *et al.* (*APSR*, 1998)



- Popular in social psychology

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)

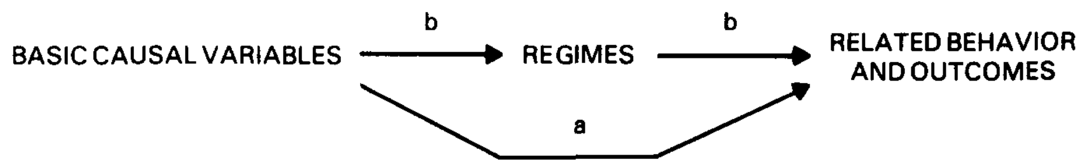


Figure 2

- Power and interests are mediated by regimes

Current Practice in the Discipline

- Regression

$$Y_i = \alpha + \beta T_i + \gamma M_i + \delta X_i + \epsilon_i$$

- Each coefficient is interpreted as a causal effect
- Sometimes, it's called marginal effect
- Idea: increase T_i by one unit while holding M_i and X_i constant

- The Problem: Post-treatment bias
- If you change T_i, that may also change M_i
- Usual advice: only include causally prior (or pre-treatment) variables
- But, then you lose causal mechanisms!

Statistical Framework of Causal Inference

- Units: $i = 1, \ldots, n$
- "Treatment": $T_i = 1$ if treated, $T_i = 0$ otherwise
- *Observed* outcome: Y_i
- Pre-treatment covariates: X_i
- Potential outcomes: $Y_i(1)$ and $Y_i(0)$ where $Y_i = Y_i(T_i)$

Voters	Contact	Turnout		Age	Party ID
i	T_i	$Y_i(1)$	$Y_i(0)$	X_i	X_i
1	1	1	?	20	D
2	0	?	0	55	R
3	0	?	1	40	R
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	1	0	?	62	D

- Causal effect: $Y_i(1) - Y_i(0)$

Notation for Causal Mediation Analysis

- Binary treatment (can be generalized): $T_i \in \{0, 1\}$
- Mediator: M_i
- Outcome: Y_i
- Observed covariates: X_i

- Potential mediators: $M_i(t)$ where $M_i = M_i(T_i)$
- Potential outcomes: $Y_i(t, m)$ where $Y_i = Y_i(T_i, M_i(T_i))$

Defining and Interpreting Causal Mediation Effects

- Total causal effect:

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

- Causal mediation effects:

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

- Change the mediator from $M_i(0)$ to $M_i(1)$ while holding the treatment constant at t
- Indirect effect of the treatment on the outcome through the mediator under treatment status t
- $Y_i(t, M_i(t))$ is observable but $Y_i(t, M_i(1-t))$ is not

- Different from *controlled* direct effects: $Y_i(t, m) - Y_i(t, m')$
- Not applicable if the mediator is manipulated

- Direct effects:

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

- Change the treatment from 0 to 1 while holding the mediator constant at $M_i(t)$

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

$$\tau_i \;=\; \delta_i(t) + \zeta_i(1-t) \;=\; \frac{1}{2}\sum_{t=0}^{1}\{\delta_i(t) + \zeta_i(t)\}$$

- Quantities of interest: Average Causal Mediation Effects,

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

The Proposed Identification Assumption

Assumption 1 (Sequential Ignorability)

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

- $\{Y_i(t, m), M_i(t)\} \perp\!\!\!\perp T_i = t \mid X_i = x$ is not sufficient
- $Y_i(t, m) \perp\!\!\!\perp M_i \mid T_i = t, X_i = x$ is not sufficient

- Weaker than Pearl (2001) if the treatment is randomized
- Cannot condition on post-treatment confounders that are causally prior to the mediator
- If such confounders are exist, an additional assumption, e.g., no-interaction assumption, is necessary (Robins)

Nonparametric Identification and Inference

Theorem 1 (Nonparametric Identification)

Under Assumption 1,

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

- Two regressions:

$$\mu_{tm}(x) \equiv \mathbb{E}(Y_i \mid T_i = t, M_i = m, X_i = x),$$
$$\lambda_t(x) \equiv f(M_i \mid T_i = t, X_i = x).$$

- When M_i is discrete, $\lambda_{tm}(x) \equiv \Pr(M_i = m \mid T_i = t, X_i = x)$, and

$$\hat{\delta}(t) = \frac{1}{n}\left\{\sum_{i=1}^{n}\sum_{m=0}^{J-1} \hat{\mu}_{tm}(X_i)\left(\hat{\lambda}_{1m}(X_i) - \hat{\lambda}_{0m}(X_i)\right)\right\}.$$

Linear Structural Equation Model

Theorem 2 (Identification under LSEM)

Consider the following linear structural equation model

$$\begin{aligned} M_i &= \alpha_2 + \beta_2 T_i + \epsilon_{2i}, \\ Y_i &= \alpha_3 + \beta_3 T_i + \gamma M_i + \epsilon_{3i}. \end{aligned}$$

Under Assumption 1, the average causal mediation effects are identified as $\bar{\delta}(0) = \bar{\delta}(1) = \beta_2\gamma$.

- Run two regressions and multiply two coefficients (Baron-Kenny)!
- No need to run: $Y_i = \alpha_1 + \beta_1 T_i + \epsilon_{1i}$
- Direct effect: β_3
- Total effect: $\beta_2\gamma + \beta_3$

- Relaxing the no-interaction assumption:

$$Y_i = \alpha_3 + \beta_3 T_i + \gamma M_i + \kappa T_i M_i + \epsilon_{2i}$$

- Then, $\bar{\delta}(t) = \beta_2(\gamma + t\kappa)$
- The product formula applies to the nonparametric identification with a binary mediator

$$\begin{aligned} \bar{\delta}(t) = & \{\mathbb{E}(Y_i \mid M_i = 1, T_i = t, X_i) - \mathbb{E}(Y_i \mid M_i = 0, T_i = t, X_i)\} \\ & \times \{\Pr(M_i = 1 \mid T_i = 1, X_i) - \Pr(M_i = 1 \mid T_i = 0, X_i)\} \end{aligned}$$

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?

- Parametric 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

Parametric 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

Theorem 3 (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 $\sigma_j^2 \equiv \mathrm{var}(\epsilon_{ji})$ *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!)

Facilitating Interpretation

- How big is ρ?
- An unobserved (pre-treatment) confounder formulation:

$$\epsilon_{2i} \;=\; \lambda_2 U_i + \epsilon'_{2i} \quad \text{and} \quad \epsilon_{3i} \;=\; \lambda_3 U_i + \epsilon'_{3i},$$

- Assume $Y_i(t', m) \perp\!\!\!\perp M_i \mid T_i = t, U_i = u$
- Assume also $\epsilon'_{2i} \perp\!\!\!\perp U_i$ and $\epsilon'_{3i} \perp\!\!\!\perp U_i$
- Proportion of previously unexplained variance explained by the unobserved confounder

$$R_M^{2*} \equiv \frac{\text{var}(\epsilon_{2i}) - \text{var}(\epsilon'_{2i})}{\text{var}(\epsilon_{2i})} \quad \text{and} \quad R_Y^{2*} \equiv \frac{\text{var}(\epsilon_{3i}) - \text{var}(\epsilon'_{3i})}{\text{var}(\epsilon_{3i})}$$

- Proportion of original variance explained by the unobserved confounder

$$\widetilde{R}_M^2 \equiv \frac{\text{var}(\epsilon_{2i}) - \text{var}(\epsilon'_{2i})}{\text{var}(M_i)} \quad \text{and} \quad \widetilde{R}_Y^2 \equiv \frac{\text{var}(\epsilon_{3i}) - \text{var}(\epsilon'_{3i})}{\text{var}(Y_i)}$$

- Specify $\text{sgn}(\lambda_2\lambda_2)$ and ${R_M^*}^2, {R_Y^*}^2$ (or $\widetilde{R}_M^2, \widetilde{R}_Y^2$)

$$\rho \;=\; \text{sgn}(\lambda_2\lambda_3) R_M^* R_Y^* \;=\; \frac{\text{sgn}(\lambda_2\lambda_3)\widetilde{R}_M \widetilde{R}_Y}{\sqrt{(1 - R_M^2)(1 - R_Y^2)}},$$

where R_M^2 and R_Y^2 are based on

$$M_i \;=\; \alpha_2 + \beta_2 T_i + \epsilon_{2i}$$
$$Y_i \;=\; \alpha_3 + \beta_3 T_i + \gamma M_i + \epsilon_{3i}$$

Political Psychology Experiment: Nelson *et al.* (*APSR*)

- How does media framing affect citizens' political opinions?
- News stories about the Ku Klux Klan rally in Ohio
- Free speech frame ($T_i = 0$) and public order frame ($T_i = 1$)
- Randomized experiment with the sample size $= 136$

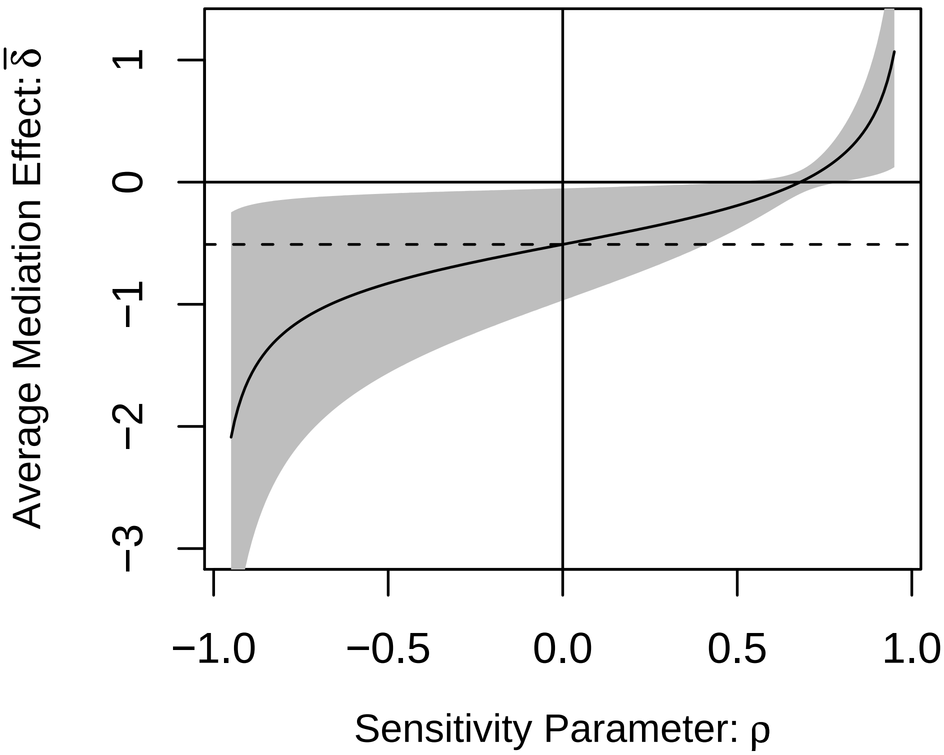
- Mediator: a scale measuring general attitudes about the importance of public order
- Outcome: a scale measuring tolerance for the Klan rally
- Expected findings: negative mediation effects

Analysis under Sequential Ignorability

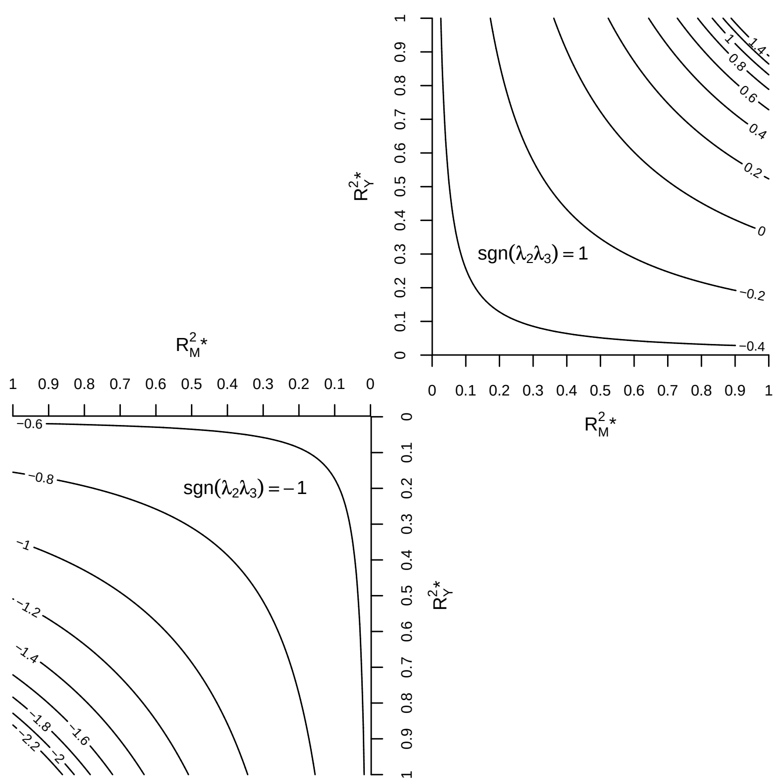
	Parametric	Nonparametric
Average Mediation Effects		
Free speech frame $\hat{\delta}(0)$	-0.451	-0.374
	$[-0.871, -0.031]$	$[-0.823, 0.074]$
Public order frame $\hat{\delta}(1)$	-0.566	-0.596
	$[-1.081, -0.050]$	$[-1.168, -0.024]$
Average Total Effect $\hat{\tau}$	-0.540	-0.627
	$[-1.207, 0.127]$	$[-1.153, -0.099]$
With the no-interaction assumption		
Average Mediation Effect	-0.510	
$\hat{\delta}(0) = \hat{\delta}(1)$	$[-0.969, -0.051]$	
Average Total Effect $\hat{\tau}$	-0.540	
	$[-1.206, 0.126]$	

Parametric Sensitivity Analysis

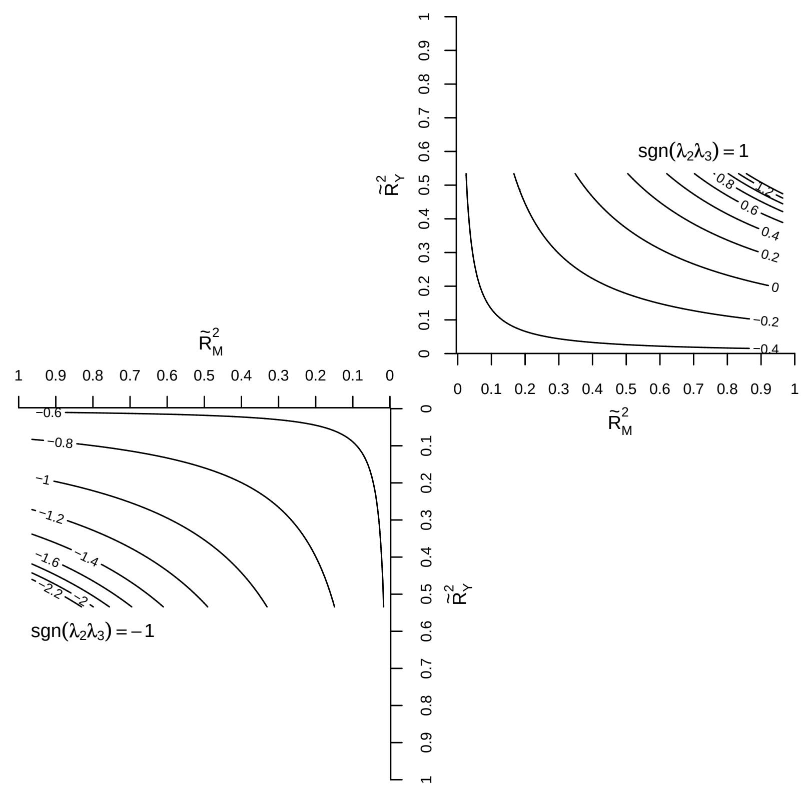
- Unobserved pre-treatment confounder (e.g., political ideology)



Proportion of unexplained variance explained by an unobserved confounder



Proportion of original variance explained by an unobserved confounder



Concluding Remarks and Work in Progress

- Quantitative analysis can be used to identify causal mechanisms!
- Estimate causal mediation effects rather than marginal effects
- Wide applications in social science disciplines

- Generalization: identification, inference, and sensitivity analysis
- linear and nonlinear relationships
- parametric and nonparametric models
- continuous and discrete mediators
- various outcome data types
- multiple mediators
- development of easy-to-use statistical software