A General Approach to Causal Mediation Analysis

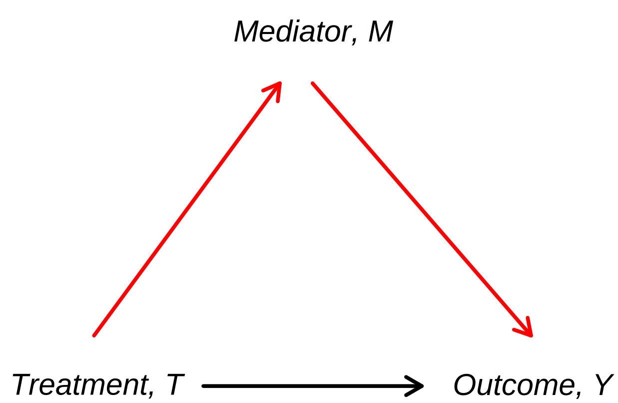
Kosuke Imai

Princeton University

August 6, 2009
Joint work with Luke Keele, Dustin Tingley, and Teppei Yamamoto

Causal Mediation Analysis

- Graphical representation



- Quantities of interest: Direct and indirect effects
- Fast growing methodological literature

Notation for Causal Mediation Analysis

- Binary treatment: $T_i \in \{0, 1\}$
- Mediator: $M_i \in \mathcal{M}$
- Outcome: $Y_i \in \mathcal{Y}$
- Observed covariates: $X_i \in \mathcal{X}$

- 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:

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

- Causal effect of the change in M_i on Y_i that would be induced by treatment
- Change the mediator from $M_i(0)$ to $M_i(1)$ while holding the treatment constant at t

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

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

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

Nonparametric Identification

- Quantity of Interest: Average Causal Mediation Effects

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

- Problem: $Y_i(t, M_i(t))$ is observed but $Y_i(t, M_i(1-t))$ can *never* be observed

- Proposed identification assumption: Sequential Ignorability

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

Theorem 1 (Nonparametric Identification)

Under sequential ignorability,

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

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

Inference Under Sequential Ignorability

- Model outcome and mediator
- Outcome model: $p(Y_i \mid T_i, M_i, X_i)$
- Mediator model: $p(M_i \mid T_i, X_i)$
- A simplest setup: Linear Structural Equation Model (LSEM)

$$
\begin{aligned}
M_i &= \alpha_2 + \beta_2 T_i + \epsilon_{i2}, \\
Y_i &= \alpha_3 + \beta_3 T_i + \gamma M_i + \epsilon_{i3}.
\end{aligned}
$$

Theorem 2 (Identification Under LSEM)

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

- Can include the interaction between T_i and M_i
- Can use parametric or nonparametric regressions; probit, logit, ordered mediator, GAM, quantile regression, etc.

Algorithm for Estimating Causal Mediation Effects

Parametric models: Quasi-Bayesian approximation

- Step 1: Estimate models for outcome and mediator
- Step 2: Take J draws from the asymptotic sampling distribution of model parameters
- Step 3: Repeat the following steps for each $j = 1, 2, ..., J$
  1. Sample K copies of $M_i(t)$ from the mediator model
  2. Given this draw, sample one copy of $Y_i(t', M_i(t))$ from the outcome model
  3. Compute QoI based on these K sets of draws
- Step 4: Compute the point estimate and uncertainty estimates from the resulting J draws of QoI

Nonparametric/semiparametric models: Nonparametric bootstrap

General implementation for statistical software

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: $\rho \equiv \mathrm{Corr}(\epsilon_{i2}, \epsilon_{i3})$
- Sequential ignorability implies $\rho = 0$
- Set ρ to different values and see how mediation effects change

Theorem 3

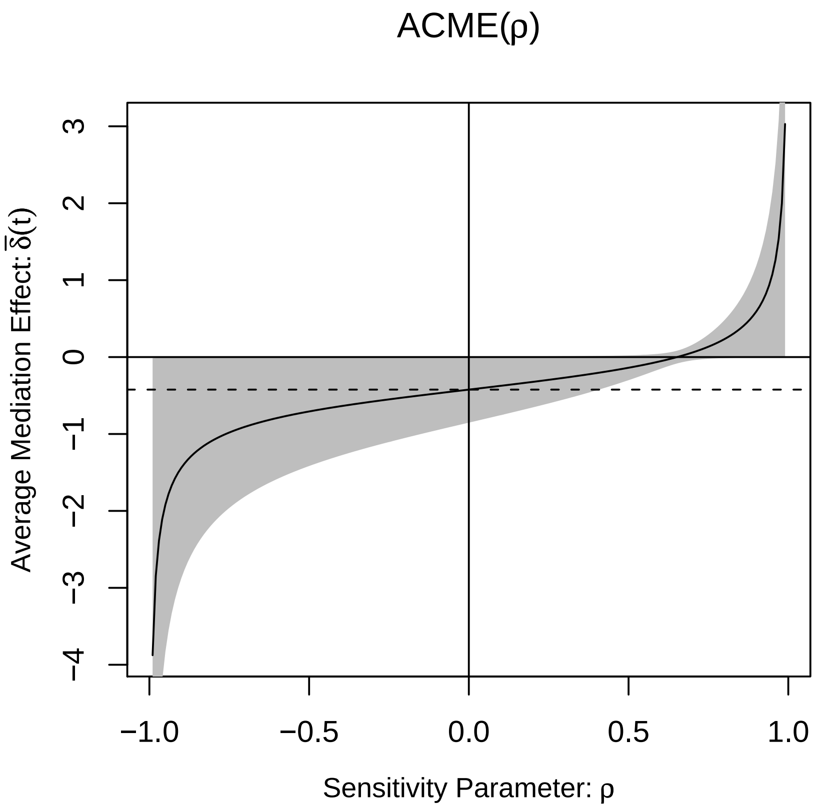
$$\bar{\delta}(0) = \bar{\delta}(1) = \frac{\beta_2 \sigma_1}{\sigma_2}\left\{\tilde{\rho} - \rho\sqrt{(1-\tilde{\rho}^2)/(1-\rho^2)}\right\},$$

where $\sigma_j^2 \equiv \mathrm{var}(\epsilon_{ij})$ *for* $j = 1, 2$ *and* $\tilde{\rho} \equiv \mathrm{Corr}(\epsilon_{i1}, \epsilon_{i2})$.

- When do my results go away completely?
- $\bar{\delta}(t) = 0$ if and only if $\rho = \tilde{\rho}$
- Easy to estimate from the regression of Y_i on T_i:

$$Y_i = \alpha_1 + \beta_1 T_i + \epsilon_{i1}$$

Sensitivity Analysis with Respect to ρ



Interpreting Sensitivity Analysis with R squares

- Interpreting ρ: how small is too small?
- An unobserved (pre-treatment) confounder formulation:

$$\epsilon_{i2} = \lambda_2 U_i + \epsilon'_{i2} \quad \text{and} \quad \epsilon_{i3} = \lambda_3 U_i + \epsilon'_{i3}$$

- How much does U_i have to explain for our results to go away?
- Sensitivity parameters: R squares
  1. Proportion of previously unexplained variance explained by U_i

$$R_M^{2*} \equiv 1 - \frac{\text{var}(\epsilon'_{i2})}{\text{var}(\epsilon_{i2})} \quad \text{and} \quad R_Y^{2*} \equiv 1 - \frac{\text{var}(\epsilon'_{i3})}{\text{var}(\epsilon_{i3})}$$

  2. Proportion of original variance explained by U_i

$$\widetilde{R}_M^2 \equiv \frac{\text{var}(\epsilon_{i2}) - \text{var}(\epsilon'_{i2})}{\text{var}(M_i)} \quad \text{and} \quad \widetilde{R}_Y^2 \equiv \frac{\text{var}(\epsilon_{i3}) - \text{var}(\epsilon'_{i3})}{\text{var}(Y_i)}$$

- Then reparameterize ρ using (R_M^{2*}, R_Y^{2*}) (or $(\widetilde{R}_M^2, \widetilde{R}_Y^2)$):

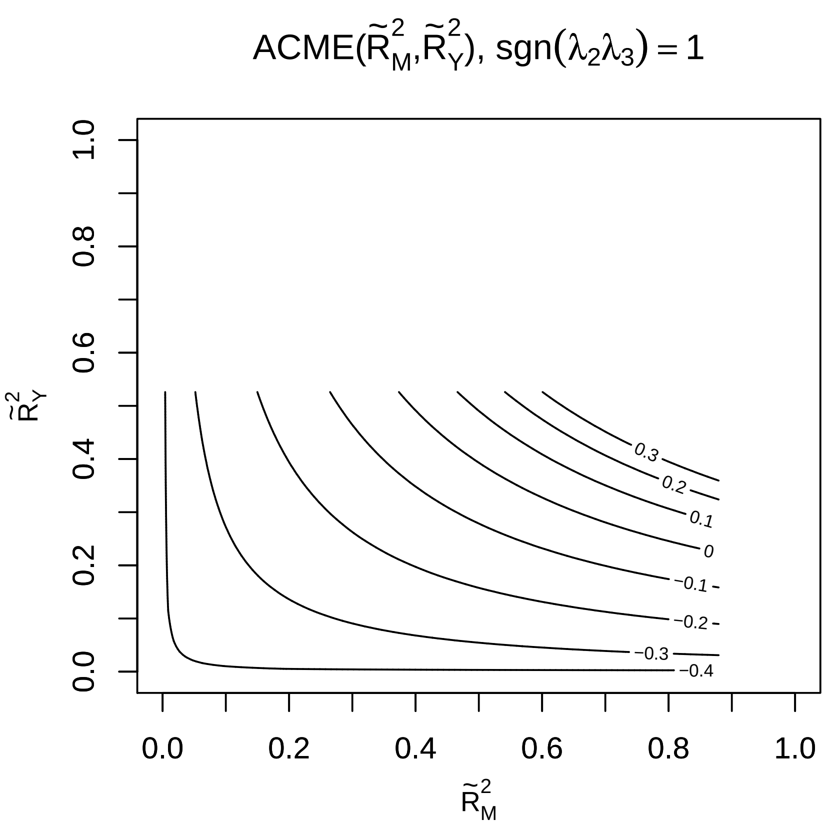
$$\rho = \text{sgn}(\lambda_2\lambda_3) R_M^* R_Y^* = \frac{\text{sgn}(\lambda_2\lambda_3)\widetilde{R}_M \widetilde{R}_Y}{\sqrt{(1 - R_M^2)(1 - R_Y^2)}},$$

  where R_M^2 and R_Y^2 are from the original mediator and outcome models

- $\text{sgn}(\lambda_2\lambda_3)$ indicates the direction of the effects of U_i on Y_i and M_i

- Set (R_M^{2*}, R_Y^{2*}) (or $(\widetilde{R}_M^2, \widetilde{R}_Y^2)$) to different values and see how mediation effects change

Sensitivity Analysis with Respect to $(\widetilde{R}_M^2, \widetilde{R}_Y^2)$



Papers and Software

- "Identification, Inference, and Sensitivity Analysis for Causal Mediation Effects."
- "A General Approach to Causal Mediation Analysis."
- "Causal Mediation Analysis in R."
- All available at
  http://imai.princeton.edu/projects/mechanisms.html

- **mediation**: R package for causal mediation analysis
- Available at
  http://cran.r-project.org/web/packages/mediation/