THE QUESTION OF MEDIATION
(direct vs. indirect effects)

1. Why decompose effects?
2. What is the definition of direct and indirect effects?
3. What are the policy implications of direct and indirect effects?
4. When can direct and indirect effect be estimated consistently from experimental and nonexperimental data?

LEGAL IMPLICATIONS OF DIRECT EFFECT

Can data prove an employer guilty of hiring discrimination?

\[ \text{Adjust for } M? \text{ No! No!} \]

\[ \text{CDE identification is completely solved (Tian et al. 2002)} \]

DEFINITION OF INDIRECT EFFECTS

\[ \text{Indirect Effect of } X \text{ on } Y: \ IE(x_0, x_1; Y) \]

The expected change in \( Y \) when we keep \( X \) constant, say at \( x_0 \), and let \( M \) change to whatever value it would have attained had \( X \) changed to \( x_1 \).

\[ E[Y_{x_0}M_{x_1} - Y_{x_0}] \]

In linear models, \( IE = TE - DE \)
POLICY IMPLICATIONS OF INDIRECT EFFECTS

What is the indirect effect of X on Y?
The effect of Gender on Hiring if sex discrimination is eliminated.

GENDER X M QUALIFICATION
IGNORE f
Y HIRING

Deactivating a link – a new type of intervention

THE MEDIATION FORMULAS IN UNCONFOUNDED MODELS

\[ m = f(x, u_1) \]
\[ y = g(x, m, u_2) \]
\[ u_1 \text{ independent of } u_2 \]

\[ DE = \sum \mathbb{E}(Y \mid x_1) - \mathbb{E}(Y \mid x_0, m) | P(m \mid x_0) \]
\[ IE = \sum \mathbb{E}(Y \mid x_0, m) [P(m \mid x_1) - P(m \mid x_0)] \]

\[ TE = \mathbb{E}(Y \mid x_1) - \mathbb{E}(Y \mid x_0) \quad TE \neq DE + IE \]
\[ IE = \text{Fraction of responses explained by mediation (sufficient)} \]
\[ TE - DE = \text{Fraction of responses owed to mediation (necessary)} \]

NONPARAMETRIC IDENTIFICATION
(Natural mediation is a solved problem)

- The nonparametric estimability of natural (and controlled) direct and indirect effects can be determined in polynomial time given any causal graph \( G \) with both measured and unmeasured variables.
- If NDE (or NIE) is estimable, then its estimand can be derived in polynomial time.
- The algorithm is complete and was extended to any path-specific effects (Shpitser, 2013).

WHEN CAN WE IDENTIFY MEDIATED EFFECTS?

\[ y = \beta_1 m + \beta_2 x + \beta_3 x m + \beta_4 w + u_1 \]
\[ m = \gamma_1 x + \gamma_2 w + u_2 \]
\[ w = \alpha x + u_3 \]

WHAT CAN MEDIATION FORMULA DO FOR PARAMETRIC ANALYSTS?

- Multi-mediators non-linear model
- What combination of parameters gives the effect mediated by M?
- \( IE(M) = \beta_1 (\gamma_1 + \alpha \gamma_2) \)
- What combination of parameters gives the effect owed to \( M \)?
- \( TE - DE(M) = (\beta_1 + \beta_3) (\gamma_1 + \alpha \gamma_2) \)
FRIENDLY EXCHANGE CONCERNING IGNORABILITY VS. GRAPHICAL ASSUMPTIONS

- Psychological Methods (2014) Imai et al. proved that graphical and ignorability assumptions are identical for randomized treatments.
- Consensus achieved regarding transparency of graphical assumptions.
- Semi-consensus regarding other aspects of the graphical vs. ignorability languages.

FORMULATING A PROBLEM IN THREE LANGUAGES

1. English: Smoking ($X$), Cancer ($Y$), Tar ($Z$), Genotypes ($U$)

2. Counterfactuals: 
   \[ Z_x(u) = Z_{xy}(u) \]
   \[ X_z(u) = X_{zy}(u) = X_z(u) = X(u) \]
   \[ Y_z(u) = Y_{xy}(u), \quad Z_z \perp \{Y_z, X\} \]

   Not too friendly:
   Consistent?, complete?, redundant?, plausible?, testable?

DAGS VS. POTENTIAL OUTCOMES: AN UNBIASED PERSPECTIVE

1. Semantic Equivalence

2. Both are abstractions of Structural Equation Models (SEM).

   \[ Y_{x}(u) = Y_{M_{x}}(u) \]
   \[ X \rightarrow Y \]
   \[ y = f(x, z, u) \]

   \[ Y_{f(u)} = \text{All factors that affect } Y \text{ when } X \text{ is held constant at } X=x. \]

FORMULATING A PROBLEM IN THREE LANGUAGES

1. English: Smoking ($X$), Cancer ($Y$), Tar ($Z$), Genotypes ($U$)

2. Counterfactuals: 
   \[ Z_x(u) = Z_{xy}(u) \]
   \[ X_z(u) = X_{zy}(u) = X_z(u) = X(u) \]
   \[ Y_z(u) = Y_{xy}(u), \quad Z_z \perp \{Y_z, X\} \]

3. Structural:
   \[ X \rightarrow Z \rightarrow Y \]
   \[ z = f_2(x, \varepsilon_2) \quad \varepsilon_1 \perp \varepsilon_2 \perp \varepsilon_3 \]
THE STRUCTURAL-COUNTERFACTUAL SYMBIOSIS

1. Express theoretical assumptions in structural language.
2. Express queries in counterfactual language.
3. Translate (1) into (2) for algebraic analysis, or (2) into (1) for graphical analysis.
4. Use either graphical or algebraic machinery to answer the query in (2).

THE MEDIATION FALLACY OR HOW TRADITIONALISTS CONFUSED AN ENTIRE CENTURY

A – Standard model
DE = Whatever changes we see in Y when we vary X and “hold M constant”
• Holding M constant ≠ controlling for M
• Statistics has no operator for “holding M constant”
• “Controlling for M” leads to fallacies in the presence of unobserved confounders

THE MEDIATION FALLACY OR HOW TRADITIONALISTS CONFUSED AN ENTIRE CENTURY

B – Conditioning on M creates dependency between T and Y though DE = 0
C – “Fixing M” correctly shows DE = 0 ⇒ No dependence

Thank you

JON’S QUESTIONS TO PANEL

1. Do you think an experiment has any value without mediational analysis?
2. Is a separate study directly manipulating the mediator useful?
3. Imai’s correlated residuals test seems valuable for distinguishing fake from genuine mediation. Is that so?
4. Why isn’t it easy to test whether participants who show the largest increases in the posited mediator show the largest changes in the outcome?
5. Why is mediational analysis any “worse” than any other method of investigation?

CHRISTIAN’S QUESTIONS TO PANEL

1. Can we go beyond “if assumptions, then conclusions”?
   • Yes, testable implications, experimental evidence, other studies.
2. How your framework would use the results of one mediation analysis to inform the setup of a second, new mediation analysis?
   • It is not a question of “framework” but of “information”
DIVIDE AND CONQUER MAKES A DIFFERENCE

Two separate sets can accomplish what their union can’t.

W_2 \text{ deconfounds } X \rightarrow M
W_3 \text{ deconfounds } X \rightarrow Y

DIVIDE AND CONQUER MAKES A DIFFERENCE

\{W_2, W_3\} \text{ confounds } X \rightarrow M

Two separate sets can accomplish what their union can’t.
To identify NDE, we must condition on W_2 and W_3 separately.

SEQUENTIAL IGNORABILITY IS NOT NEEDED

W = \emptyset

No set can deconfound X \rightarrow M.
Measuring Z permits the identification of
\( P(M = m | \text{do}(x)) \) through the front-door formula.
NDE is identified.