Modeling Mediation: Causes, Markers, and Mechanisms

Stephen W. Raudenbush
University of Chicago

Address at the Society for Research on Educational Effectiveness, Washington, DC, March 3, 2011. Many thanks to Guanglei Hong and Sean Reardon for sharing their ideas for this talk. Research reported here was supported by the W.T. Grant Foundation for a project entitled “Building Capacity to Support Group Randomized Experiments.”
Sea Change in Ed Research!

• Role of Experiments
  – Cluster randomized trials
  – Smart designs, Well powered!

• Sound Quasi-Experiments
  – Regression Discontinuity
  – Propensity score adjustment
  – Instrumental variables
What Experiments Can Tell Us

• Clear answer to sharp causal questions:
  – Average effect of assignment to treatments
  – If high compliance, average effect of an intervention
  – For whom does the intervention work?
  – In what settings does it work?
What about “How?” and “Why?”

• Example: IES study of pre-school curricula
  – Tells us which, on average, work better
  – But why
    • More time of academic instruction?
    • Better content?
    • Increase in teacher skill?
    • Compensate for lack of teacher skill?
    • Heighten child engagement?
Questions of Mediation

• In lab experiments, conduct a subsequent “factorial” randomized study
  – E.g., curricula by instructional time

• Very costly in large-scale field studies
Major Progress in Measurement
Some Major Progress

• We can ask
  – Was the treatment implemented
  – Did the treatment assignment affect practice?
  – Did the treatment increase student engagement

• These help interpret average causal effects
More challenging

• Did the treatment affect $Y$ by
  
  – Increasing teacher knowledge?
  – Increasing teacher skill?
  – Improving class climate?
  – Increasing student engagement?

• These are questions of mediation
Models for Mediation

• Conventional methods under scrutiny
• Alternatives
  – Natural direct and indirect effects
  – Controlled direct effects
  – Direct Effects within Principal Strata
  – Instrumental variables for indirect effects
• Each requires assumptions
• Methods of Estimation not widely accessible
Punch Line

- Each approach answers a different question
- So let’s think hard about our questions
Organization of Talk

Consider questions about

- Demographic markers that predict access to treatments that predict outcomes
- Treatments that predict access to future treatments that predict outcomes
- Treatments that affect surrogate markers that predict outcomes

In light of these questions consider

- Oaxaca Decomposition
- Natural Direct Effects
- Controlled Direct Effects
- Principal Stratification
- Instrumental Variables

Reflect on the larger project: finding mechanisms
“All-purpose” mediation model
(Blau and Duncan, 1965; Baron and Kenny and Baron 1986)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Mediator</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic marker</td>
<td>Treatment (years of ed.)</td>
<td>Outcome (earnings)</td>
</tr>
<tr>
<td>(parent SES)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment (new curriculum v old)</td>
<td>Mediating treatment (quality of instruction)</td>
<td>Outcome (learning)</td>
</tr>
<tr>
<td>Treatment (pre-school program)</td>
<td>Surrogate marker (early vocab)</td>
<td>Outcome (Reading comp)</td>
</tr>
</tbody>
</table>
Start with Case 1
Demographic Marker--->Treat--->Y

Example:

Does school mobility explain B-W test score gap? (Raudenbush, Jean, Art, 2011)
Racial disparity $\beta = \gamma \delta + \theta$
But we found differential effects!

\[ \gamma = (\phi_b - \phi_w) = \text{differential exposure} \]
\[ \delta_b - \delta_w = \text{differential vulnerability} \]

**Oaxaca decomposition:**

\[ \beta = (\phi_b - \phi_w)\delta_b \quad (\text{indirect path}) \]
\[ + \theta_{\text{low}} + \phi_w(\delta_b - \delta_w) \quad (\text{direct path}) \]
Decomposition is not unique

$$\beta = (\phi_b - \phi_w) \delta_b \quad (\text{indirect path})$$
$$+ \theta_{\text{low}} + \phi_w (\delta_b - \delta_w) \quad (\text{direct path})$$

$$= (\phi_b - \phi_w) \delta_w \quad (\text{indirect path})$$
$$+ \theta_{\text{low}} + \phi_b (\delta_b - \delta_w) \quad (\text{direct path})$$
Alternative expression

\[
\beta = (\phi_b - \phi_w) \delta_b \quad (indirect \ path) \\
+ \phi_w \theta_{\text{high}} + (1 - \phi_w) \theta_{\text{low}} \quad (direct \ path)
\]
Case 1 is easy!

- We have to estimate only 1 causal effect for Blacks and 1 for Whites
- Assume ignorable assignment to treatment within race
- Eg: How much does Class size reduction reduce the B-W gap (Krueger and Whitmore, 19xx)
Case 2

Treatment → Mediating Treatment → Y

Example:

Does a new curriculum reduce aggression by improving the emotional climate of the classroom?

(VanderWeele, Hong, Jones, and Brown, 2011)
Assignment to 4Rs Induces Better Climate

- 4 R’s
  - 1 = Treatment
  - 0 = Control

- Emotional climate
  - 1 = good
  - 0 = bad

- Aggression
Must estimate causal effects of a sequence of two treatments!

The first (randomized) treatment increases the probability of exposure to the second (non-randomized) treatment.

Assumptions are subtle and strong!
Person-specific Potential Outcomes

\[ T = 1 \]

\[ M(1) = 1 \rightarrow Y(1,1) \]
\[ M(1) = 0 \rightarrow Y(1,0) \]

\[ T = 0 \]

\[ M(0) = 1 \rightarrow Y(0,1) \]
\[ M(0) = 0 \rightarrow Y(0,0) \]
Person-specific Causal Effects

- **Effect of** $T$ **on** $M$

  $$\Gamma = M(1) - M(0)$$

- **Effect of** $M$ **on** $T$

  $$\Delta_1 = Y(1,1) - Y(1,0) \quad \text{effect if } T = 1$$
  $$\Delta_0 = Y(0,1) - Y(0,0) \quad \text{effect if } T = 0$$

- **Direct Effect of** $T$ **on** $Y$

  $$\Theta_1 = Y(1,1) - Y(0,1) \quad \text{effect if } M = 1$$
  $$\Theta_0 = Y(1,0) - Y(0,0) \quad \text{effect if } M = 0$$
Natural Direct and Indirect Effects

- Total Effect: person-specific “Oaxaca Decomposition!”

\[ B = Y(1, M(1)) - Y(0, M(0)) \]

\[ = Y(1, M(1)) - Y(1, M(0)) \quad \text{natural indirect effect} \]
\[ + Y(1, M(0)) - Y(0, M(0)) \quad \text{natural direct effect} \]

\[ = \Gamma \Delta_1 \quad \text{natural indirect effect} \]
\[ + \Theta_0 + M(0)(\Delta_1 - \Delta_0) \quad \text{natural direct effect} \]
Person-specific Causal Model

\[ \Theta_0 + M(0)(\Delta_1 - \Delta_0) \]

Total (ITT) Effect

\[ B = \Gamma \Delta_1 + \Theta_0 + M(0)(\Delta_1 - \Delta_0) \]
But what is the _average_ causal effect?

\[ E(B) = \beta = E[\Gamma \Delta_1 + \Theta_0 + M(0)(\Delta_1 - \Delta_0)] \]

\[ = \gamma \delta_1 + \text{Cov}(\Gamma, \Delta_1) \]
\[ + \Pr(M(0) = 1)(\delta_1 - \delta_0) + \text{Cov}(M(0), \Delta_1 - \Delta_0) \]

\[ = \gamma \delta_1 + \theta_0 + \Pr(M(0) = 1)(\delta_1 - \delta_0) \]
\[ \quad \text{(if we assume away covariances)} \]

\[ = \gamma \delta + \theta \]
\[ \quad \text{(if we assume away differential effect of mediator)} \]
Summary on Natural Direct and Indirect Effects

– Assume ignorable assignment to T
– Assume ignorable assignment to M within levels of T
– Assume away troublesome covariances
– Estimate via a regression method (Peterson et al.) or a non-parametric weighting method (Hong)
– Assume in addition no differential effects and linearity and use conventional methods!
Controlled Direct Effects

- Estimate

\[ \theta_0 = E[Y(1,0) - Y(0,0)] \]
\[ \theta_1 = E[Y(1,1) - Y(0,1)] \]

- Still need ignorable assignment to T
- Ignorable assignment to M given T
- Eliminates Troublesome covariances!
- But no estimation of indirect effect
Instrumental Variables

• Recall $B = \Gamma \Delta_1$ natural indirect effect
  
  $\Theta_0 + M(0)(\Delta_1 - \Delta_0)$ natural direct effect

• If we assume direct effect=0, we have
  
  $B = \Gamma \Delta$  \hspace{1cm} ($\Delta = Y(M(1)) - Y(M(0))$)

• Population-average effect
  
  $E(B) = \beta = E(\Gamma \Delta) = \gamma \delta + Cov(\Gamma, \Delta)$

• Assume away covariance $\beta = E(\Gamma \Delta_1) = \gamma \delta_1$

• Assume $\gamma \neq 0$, we have $\beta / \gamma = \delta_1$
Summary on Instrumental Variables

- Assume ignorable assignment to $T$
- Assume no direct effect ("exclusion restriction")
- Assume away troublesome covariance
- Assume $T$ affects $M$ ($\gamma \neq 0$)
Case 3
Treatment $\rightarrow$ Surrogate marker $\rightarrow$ Y

Example:

Does a new pre-school curriculum increase self regulation in K and hence reduce adolescent crime?

Mediator “stands in” for outcome=crime
Mediator is not a treatment
Case 3

- T=Pre-school curriculum
  - 1=yes
  - 0=no

- M=Self-regulation in K
  - 1=yes
  - 0=no

- Later Crime=Y
# Principal stratification

<table>
<thead>
<tr>
<th></th>
<th>M(1)</th>
<th>M(0)</th>
<th>Total Effect</th>
<th>Direct Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compliers</strong></td>
<td>1</td>
<td>0</td>
<td>$Y(1,1)-Y(0,0)$</td>
<td>----</td>
</tr>
<tr>
<td><strong>Always-takers</strong></td>
<td>1</td>
<td>1</td>
<td>$Y(1,1)-Y(0,1)$</td>
<td>$\theta_{AT}$</td>
</tr>
<tr>
<td><strong>Never-takers</strong></td>
<td>0</td>
<td>0</td>
<td>$Y(1,0)-Y(0,0)$</td>
<td>$\theta_{NT}$</td>
</tr>
<tr>
<td><strong>Defiers</strong></td>
<td>0</td>
<td>1</td>
<td>$Y(1,0)-Y(0,1)$</td>
<td>----</td>
</tr>
</tbody>
</table>
Summary on Principal stratification

- How a child will respond to intervention is a pre-treatment characteristic of that child
- Causal effects defined only within strata
- Direct effects defined for strata of children whose M is not changed by T.
- Strata are not observed, must be modeled
- No sharp point estimate of causal effects
- Seems most plausible when M is a surrogate marker rather than a second treatment
Summary

Case 1 (Demographic marker $\rightarrow$ Treat $\rightarrow$ Y)
- Use Oaxaca decomposition

Case 2 (Treat1 $\rightarrow$ Treat 2 $\rightarrow$ Y)
- Natural direct and indirect effects
- Controlled direct effects
- Instrumental variables

Case 3 (Treat $\rightarrow$ S. Marker $\rightarrow$ Y)
- Principal stratification
Conclusion

• Each method makes pretty strong assumptions
• The conventional “SEM” makes very strong assumptions
• We should pursue:
  – The method that best suits the question
  – That makes weakest assumptions
• We should state and evaluate the assumptions
However,…

• The project of discovering mechanisms and explaining effects requires a broad, multidisciplinary research agenda
Study “one arrow at a time”

Instructional innovation

Student Engagement

Knowledge

Practice

Class Climate

 Outcome
References

Classical mediation

Oaxaca decomposition

Natural and controlled direct effects

Instrumental variables

Principal stratification

Empirical studies cited