An Introduction to Causal Mediation Analysis

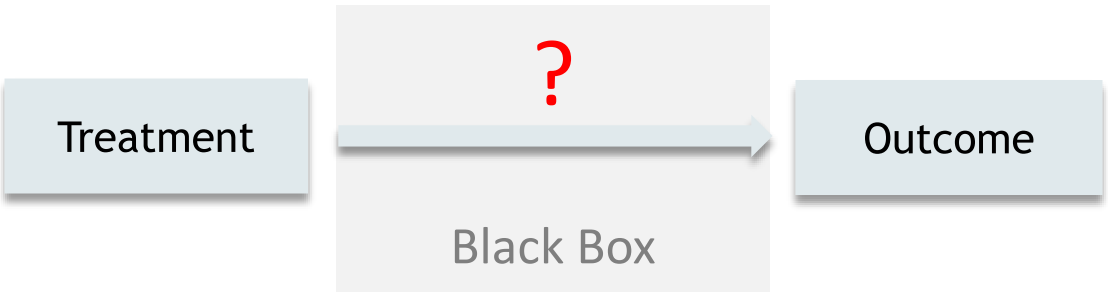
Xu Qin
University of Chicago

Presented at the Central Iowa R User Group Meetup

Aug 10, 2016

Causality

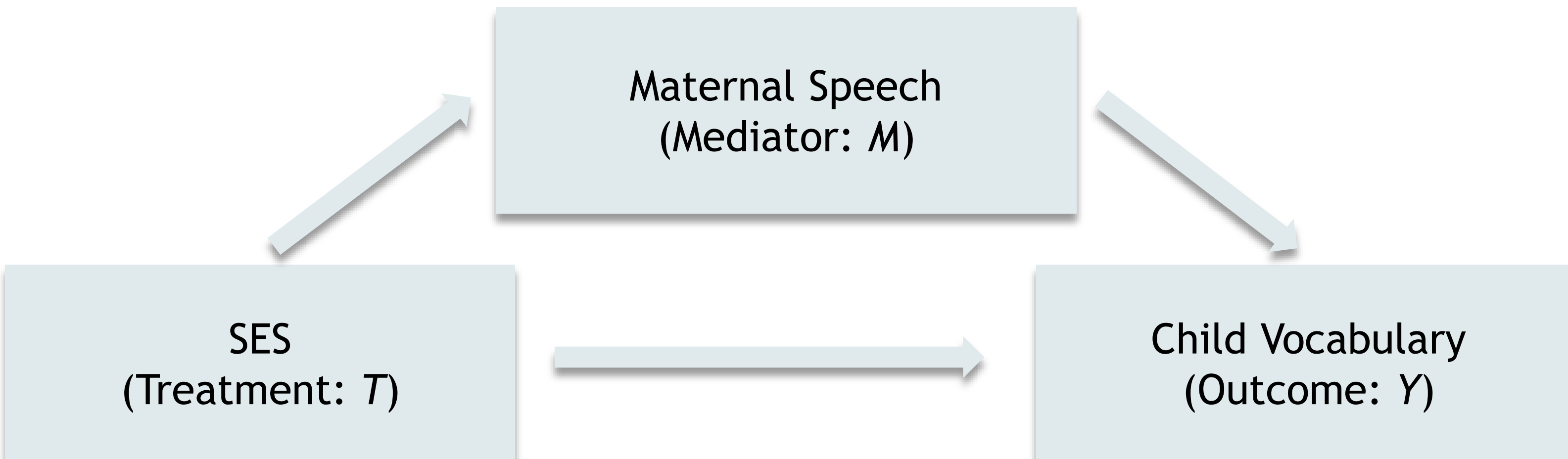
- In the applications of statistics, many central questions are related to causality rather than simply association.
  - Sociology: Does divorce affect children's education?
  - Health: Is a new drug effective against a disease?
  - Economy: Does a job training program improve participants' earnings?
  - Business: Does a sales reward program boost a company's profits?
- People care about not only the causal effect itself, but also how and why an intervention affects the outcome.



What Is Mediation?

- It uncovers the black box.
- Baron and Kenny (1986): it represents the generative mechanism through which the focal independent variable (treatment) is able to influence the dependent variable of interest (outcome).
- It decomposes the total treatment effect into an indirect effect transmitted through the hypothesized mediator and a direct effect representing the contribution of other unspecified pathways.

Example



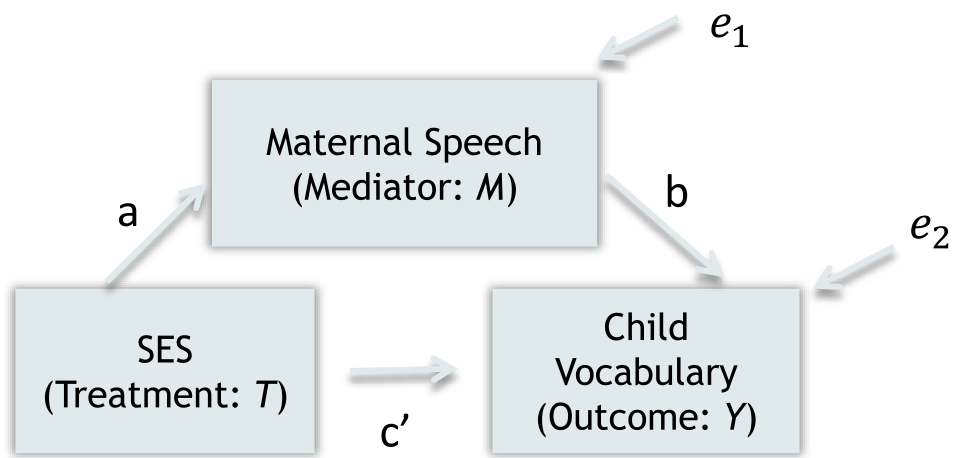
- **Direct Effect**: The impact of SES on child vocabulary without changing maternal speech.
- **Indirect Effect**: The improvement in child vocabulary attributable to the SES-induced difference in maternal speech.

Conventional Estimation Method

$$Y = d_0 + cT + e_0$$

$$M = d_1 + aT + e_1$$

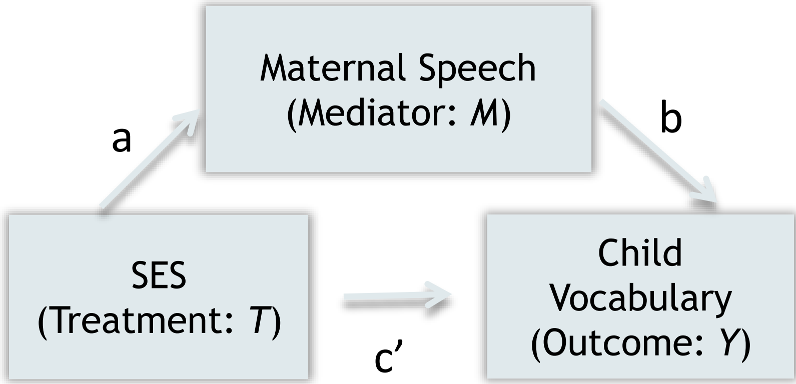
$$Y = d_2 + c'T + bM + e_2$$



(Wright, 1934; Baron and Kenny, 1986; Judd and Kenny, 1981)

- Total treatment effect: c
- Direct Effect: c'
- Indirect Effect: ab or $c - c'$
  - Significance test: Sobel test (Sobel, 1982); Bootstrapping (Bollen & Stine, 1990; Shrout & Bolger, 2002); Monte Carlo Method (MacKinnon, Lockwood, and Williams, 2004)

Limitations



- The path coefficients represent the **causal** effects of interest only when
  - the functional form of each of the models is correctly specified
  - no confounding of the T-Y relation (no covariates associated with both T and Y)
  - no confounding of the T-M relation
  - no confounding of the M-Y relation (Either pre-treatment or post-treatment)
  - no interaction exists between T and M affecting Y. However, this typically overlooks the fact that a treatment may generate an impact on the outcome through not only changing the mediator value but also changing the mediator-outcome relationship (Judd & Kenny, 1981).

Potential Outcomes Framework

Rubin (1978, 1986)

Book recommendation: Imbens and Rubin (2015)

Movie: It's a Wonderful Life
Example Source: Imbens and Rubin (2015)

POTTERSVILLE
NO PARKING
NO DOGS ALLOWED

Potential Outcomes Framework

Observed

Counterfactual

Definition of Causal Effects

Individual i's potential outcome under $T = 1$: $Y_i(1)$

Individual i's potential outcome under $T = 0$: $Y_i\ (0)$

Treatment effect for individual i: $Y_i(1) - Y_i(0)$

Population average treatment effect: $\delta \triangleq E[Y(1)] - E[Y(0)]$

ID	Treatment	Potential outcomes		Causal effect
i	T_i	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$
1	1	16	12	4
2	1	14	10	4
3	1	15	2	13
4	0	20	10	10
5	0	10	6	4
True averages		$E[Y(1)] = 15$	$E[Y(0)] = 8$	$\delta = 7$

SUTVA

- No Interference
  - The potential outcomes for any unit do not vary with the treatments assigned to other units.
- No Hidden Variations of Treatments
  - For each unit, there are no different forms or versions of each treatment level, which lead to different potential outcomes.

Identification of Causal Effects

ID	Treatment	Potential outcomes		Causal effect
i	T_i	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$
1	1	16	?	4
2	1	14	?	4
3	1	15	?	13
4	0	?	10	10
5	0	?	6	4
True averages		$E[Y(1)] = ?$	$E[Y(0)] = ?$	$\delta = ?$
Observed averages		$E[Y\|T = 1] = 15$	$E[Y\|T = 1] = 8$	$E[Y\|T = 1] - E[Y\|T = 0] = 7$

- Identification assumptions relate counterfactual quantities to observable population data. In a randomized design, ignorability assumption holds:

Under $Y_i(t) \perp\!\!\!\perp T_i$ for $t = 0,1$, $E[Y(t)] = E[Y(t)|T = t]$. Hence

$$\delta = E[Y|T = 1] - E[Y|T = 0]$$

Identification of Causal Effects

- In observational studies, we are able to identify the causal effect under strong ignorability assumption:

$$Y_i(t) \perp\!\!\!\perp T_i | \mathbf{X}_i = \mathbf{x}$$

where $0 < \mathrm{P}(T_i = 1 | \mathbf{X}_i = \mathbf{x}) < 1$

- Population average treatment effect can be identified by

$$\delta = E\{E[Y|T = 1, \mathbf{X}]\} - E\{E[Y|T = 0, \mathbf{X}]\}$$

Estimation of Causal Effects

- Propensity-score based methods (Rosenbaum and Rubin, 1983):
  - Matching
  - Subclassification
  - Covariance adjustment
  - Inverse weighting
- Sensitivity Analysis (Rosenbaum, 1986)
  - The goal is to quantify the degree to which the key identification assumption must be violated for a researcher’s original conclusion to be reversed.
- Software list (including R packages) on Prof. Elizabeth Stuart's webpage: http://www.biostat.jhsph.edu/~estuart/propensityscoresoftware.html

Causal Mediation Analysis

Book recommendation: Hong (2015); VanderWeele (2015)

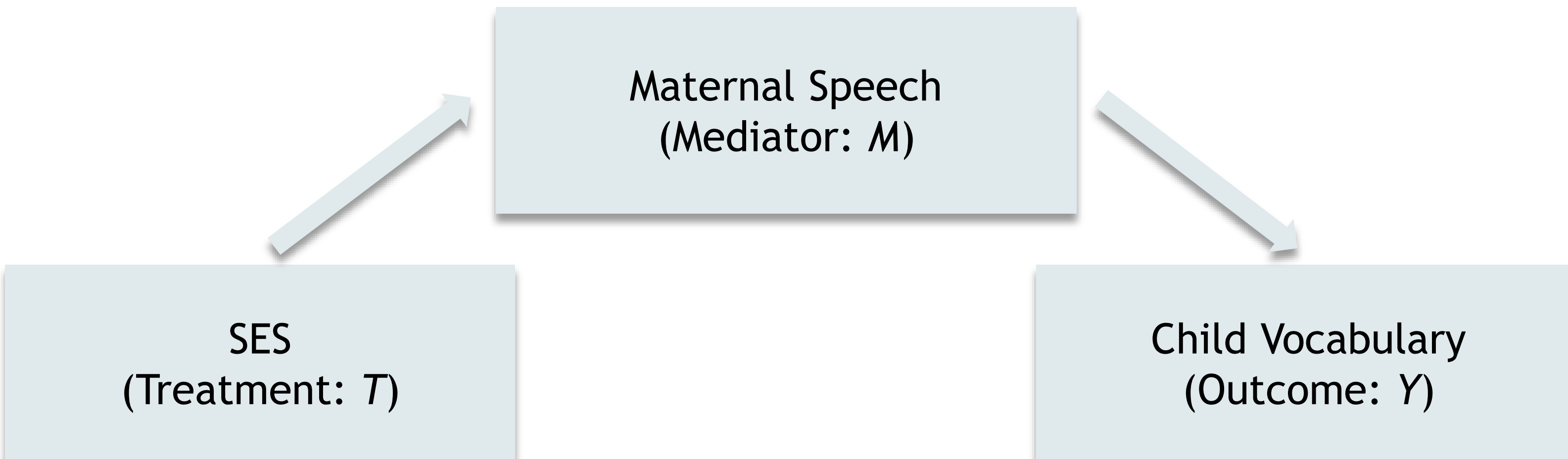
Potential Mediators and Outcomes

ID	Treatment	Potential mediators		Potential outcomes		
i	T_i	$M_i(1)$	$M_i(0)$	$Y_i\big(1, M_i(1)\big)$	$Y_i\big(0, M_i(0)\big)$	$Y_i\big(0, M_i(0)\big)$
1	1	1	1	Y(1,1)	Y(0,1)	Y(1,1)
2	1	1	0	Y(1,1)	Y(0,0)	Y(1,0)
3	1	0	0	Y(1,0)	Y(0,0)	Y(1,0)
4	0	1	1	Y(1,1)	Y(0,1)	Y(1,1)
5	0	1	0	Y(1,1)	Y(0,0)	Y(1,0)
Population Average		E[M(1)]	E[M(0)]	E[Y(1,M(1))]	E[Y(0,M(0))]	E[Y(1,M(0))]

Research Question I

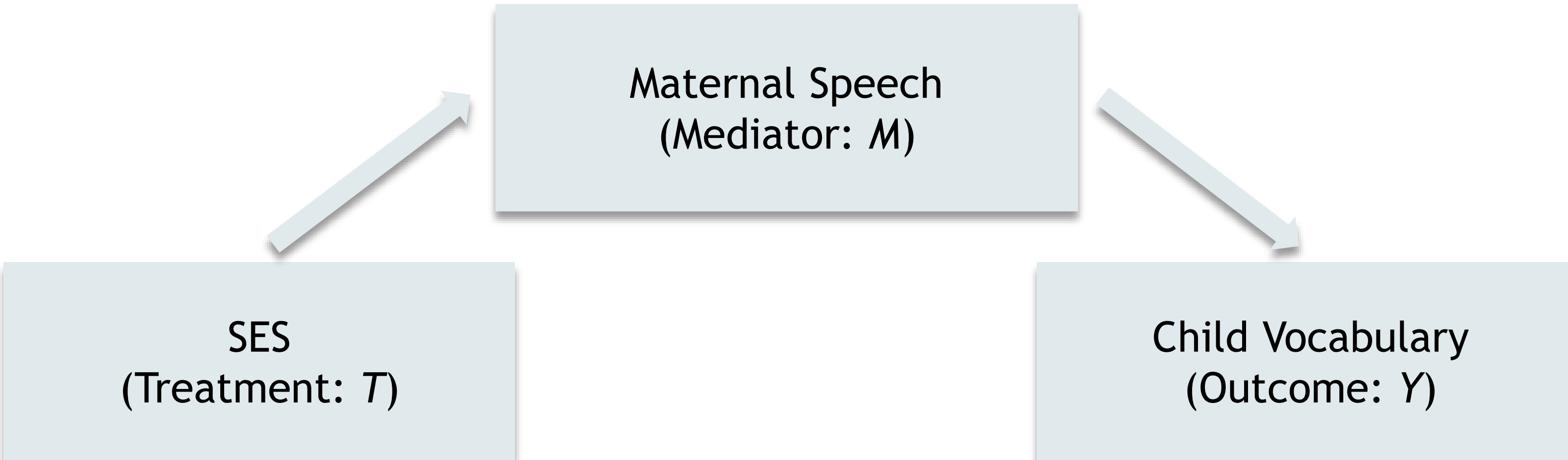
- How much is the average SES impact on child vocabulary?
- Population average total treatment effect: $E[Y(1, M(1))] - E[Y(0, M(0))]$

Research Question II



- How would the SES-induced change in maternal speech exert an impact on child vocabulary?

Definition (Pearl, 2001)



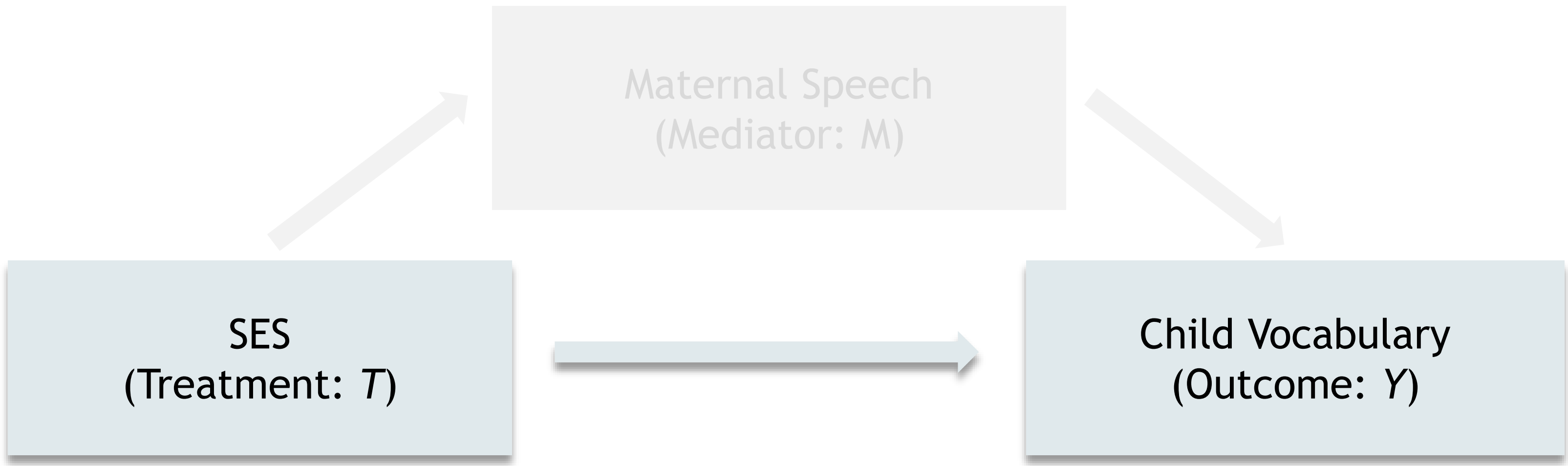
	High SES $T_i = 1$	Low SES $T_i = 0$
Maternal speech is SES is high $M_i(1)$	$\boldsymbol{Y_i(1, M_i(1))}$	
Maternal speech is SES low $M_i(0)$	$\boldsymbol{Y_i(1, M_i(0))}$	$Y_i(0, M_i(0))$

Population Average Natural Indirect Effect: $E[Y(1, M(1))] - E[Y(1, M(0))]$

Research Question III

- How much is the average causal effect of SES on child vocabulary without changing maternal speech?

Definition (Pearl, 2001)



	High SES $T_i = 1$	Low SES $T_i = 0$
Maternal speech is SES is high $M_i(1)$	$Y_i(1, M_i(1))$	
Maternal speech is SES is high $M_i(0)$	$\boldsymbol{Y_i}(\mathbf{1}, M_i(\mathbf{0}))$	$\boldsymbol{Y_i}(\mathbf{0}, M_i(\mathbf{0}))$

Population Average Natural Direct Effect: $E[Y(1, M(0))] - E[Y(0, M(0))]$

Alternative Definitions (Pearl, 2001)

- **Mechanisms**
  - Natural Indirect effect: $E\left[Y\left(1, M(1)\right)\right] - E\left[Y\left(1, M(0)\right)\right]$
  - Counterfactuals about treatment-induced mediator values
- **Manipulations**
  - Controlled direct effect: $E[Y(t, m)] - E[Y(t, m')]$
  - Causal effect of directly manipulating the mediator under T = t

Identification of Causal Effects

- We face an "identification problem" since we don't observe $Y_i(1, M_i(0))$
- Sequential Ignorability (Imai et al., 2010a, 2010b)

$$\{Y_i(t', m), M_i(t)\} \perp\!\!\!\perp T_i | \mathbf{X}_i = \mathbf{x}$$

$$Y_i(t', m) \perp\!\!\!\perp M_i(t) | T_i = t, \mathbf{X}_i = \mathbf{x}, \text{ for } t', t = 0,1$$

where $0 < \Pr(T_i = t | \mathbf{X}_i = \mathbf{x}) < 1, 0 < \Pr(M_i(t) = m | T_i = t, \mathbf{X}_i = \mathbf{x}) < 1$

  - Within levels of pretreatment confounders, the treatment is ignorable.
  - Within levels of pretreatment confounders, the mediator is ignorable given the observed treatment.

Existing Analytic Methods

- Instrumental Variable Method (Angrist, Imbens and Rubin, 1996)
  - Exclusion restriction: a constant zero direct effect
  - Assumes no T-by-M interaction
- Marginal Structural Model
  - For controlled direct effect: Robins, Hernan, and Brumback (2000)
  - For natural direct and indirect effects: VanderWeele (2009)
  - Assumes no T-by-M interaction
- Modified Regression Approach (Valeri & VanderWeele, 2013)

$$M = d_1 + aT + \beta_1 \boldsymbol{X} + e_1$$

$$Y = d_2 + c'T + bM + dTM + \beta_2 \boldsymbol{X} + e_2$$

- [Resampling Method]
- [Weighting Method]

Resampling Method (Imai et al., 2010a, 2010b)

- Algorithm 1 (Parametric)
  - Step 1: Fit models for the observed outcome and mediator variables.
  - Step 2: Simulate model parameters from their sampling distribution.
  - Step 3: Repeat the following three steps for each draw of model parameters:
    - 1. Simulate the potential values of the mediator.
    - 2. Simulate potential outcomes given the simulated values of the mediator.
    - 3. Compute quantities of interest (NDE, NIE, or average total effect).
  - Step 4: Compute summary statistics, such as point estimates (average) and confidence intervals.
  - Sensitivity analysis
- Algorithm 2 (Nonparametric/Semiparametric)
  - Combine Algorithm 1 with bootstrap
- R package: "mediation"
  - http://imai.princeton.edu/software/mediation.html
  - http://web.mit.edu/teppei/www/research/mediationR.pdf

Weighting Method

	High SES $T_i = 1$	Low SES $T_i = 0$
Maternal speech is SES is high $M_i(1)$	$E[Y(1, M(1))]$ $\parallel$ $E[Y \vert T = 1]$	
Maternal speech is SES is high $M_i(0)$	$E[Y(1, M(0))]$ $\parallel$ $E[WY \vert T = 1]$	$E[Y(0, M(0))]$ $\parallel$ $E[Y \vert T = 0]$

$$\boldsymbol{W} = \frac{\Pr(M = m \mid T = \mathbf{0}, \boldsymbol{X} = \boldsymbol{x})}{\Pr(M = m \mid T = \mathbf{1}, \boldsymbol{X} = \boldsymbol{x})}$$

Hong (2010, 2015); Hong et al. (2011, 2015); Hong and Nomi, 2012; Huber (2014); Lange et al. (2012); Lange et al. (2014); Tchetgen Tchetgen and Shpitser (2012); Tchetgen Tchetgen (2013)

Software "RMPW" could be downloaded from: hlmsoft.net/ghong

References

- Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434), 444-455
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Bollen, K. A., & Stine, R., (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, 20, 115-40.
- Hong, G. (2010). Ratio of mediator probability weighting for estimating natural direct and indirect effects. In *Proceedings of the american statistical association, biometrics section* (pp. 2401–2415). American Statistical Association.
- Hong, G. (2015). *Causality in a social world: Moderation, mediation and spill-over*. John Wiley & Sons.
- Hong, G., Deutsch, J., & Hill, H. D. (2011). Parametric and non-parametric weighting methods for estimating mediation effects: An application to the national evaluation of welfare-to-work strategies. In *Proceedings of the american statistical association, social statistics section* (pp. 3215–3229). American Statistical Association.
- Hong, G., Deutsch, J., & Hill, H. D. (2015). Ratio-of-mediator-probability weighting for causal mediation analysis in the presence of treatment-by-mediator interaction. *Journal of Educational and Behavioral Statistics*, 40 (3), 307–340.
- Hong, G., & Nomi, T. (2012). Weighting methods for assessing policy effects mediated by peer change. *Journal of Research on Educational Effectiveness*, 5 (3), 261–289.
- Huber, M. (2014). Identifying causal mechanisms (primarily) based on inverse probability weighting. *Journal of Applied Econometrics*, *29* (6), 920–943.

- Imai, K., Keele, L. and Tingley, D. (2010a) A general approach to causal mediation analysis. *Psychological methods*, 15(4), 309.
- Imai, K., Keele, L. and Yamamoto, T. (2010b) Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 25(1), 51-71.
- Imbens, G. W., & Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, 5, 602-619.
- Lange, T., Rasmussen, M., & Thygesen, L. (2014). Assessing natural direct and indirect effects through multiple pathways. *American journal of epidemiology*, *179* (4), 513.
- Lange, T., Vansteelandt, S., & Bekaert, M. (2012). A simple unified approach for estimating natural direct and indirect effects. *American journal of epidemiology*, *176* (3), 190–195.
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39, 99-128.
- Pearl, J. (2001) Direct and indirect effects. In *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*, pp. 411-420. Morgan Kaufmann Publishers Inc.
- Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 550-560.
- Rosenbaum, P. R. (1986). Dropping out of high school in the United States: An observational study. *Journal of Educational and Behavioral Statistics*, 11(3), 207-224.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70, 41–55.

- Rubin, D. B. (1978) Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics*, 6(1), 34-58.
- Rubin, D. B. (1986) Statistics and causal inference: Comment: Which ifs have causal answers. *Journal of the American Statistical Association*, 81(396), 961-962.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422-445.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology 1982* (pp. 290-312). Washington DC: American Sociological Association.
- Tchetgen, E. J. T., Shpitser, I., et al. (2012). Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis. *The Annals of Statistics*, *40* (3), 1816–1845.
- Tchetgen Tchetgen, E. J. (2013). Inverse odds ratio-weighted estimation for causal mediation analysis. *Statistics in medicine*, *32* (26), 4567–4580.
- Valeri, L., & VanderWeele, T. J. (2013). Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with sas and spss macros. *Psychological methods*, 18 (2), 137
- VanderWeele, T. J. (2009). Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*, 20(1), 18-26.
- VanderWeele, T. (2015*). Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.
- Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, 5, 161-215.