因果科学读书会

关于路径效应中的半参数估计
——以AIDS疗法的用药依从性效应研究为例

报告人：陆怡舟

就职于某商业银行；研究兴趣：金融风险管理及相关延拓领域

luyizhou@139.com

感谢：集智俱乐部、智源研究院、CaualAI

集智斑图 Swarma Pattern

Outline

- Preliminaries & Terminologies
- Causal model & Path-specific effect
- Semi-parametric estimators
- Simulation & PEPFAR results
- Authors & Further Readings

On the Scope of Casual Research

- 因果图分析（结合领域知识）
- 因果图发现（较少领域知识）
- 因果效应可估计性
- 观察数据
- 随机实验
- 因果效应估计
- 中介效应（路径效应）估计
- 估计量设计
- 相合性
- 鲁棒性
- 。。。

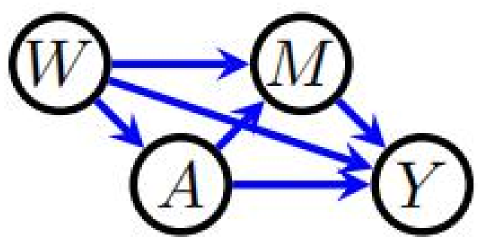
Preliminaries & Terminologies

因果效应、因果图分析：		
Potential Outcome	潜在结果	Handbook Ch15-16
Counterfactual	反事实	Handbook Ch15-16
Identifiablity	可识别（因果效应可通过观测数据识别）	上一期分享
g-formula		上一期分享
truncated factorization	截断、可分解的（条件分布）	Handbook(16.2.4, 16.2.5)
mediator-outcome confounding	关于中介变量、结果变量的混杂因子	上一期分享
recanting	关于路径效应的可识别判定标准	上一期分享

- ——“potential outcomes are also referred to as counterfactuals.” （Handbook 15.2）
- ——“allows identication of certain distributions of potential outcomes from the observed data distribution” （Handbook 16.1）

- Recursive definition of Potential outcome（16.2）

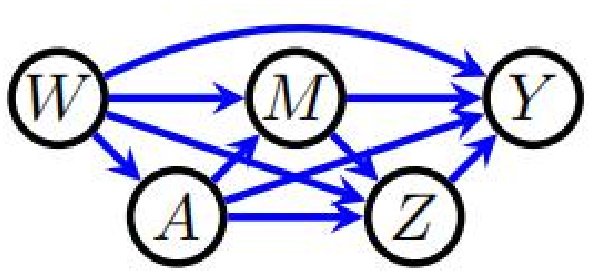
$$Y(\mathbf{a}) \equiv Y(\mathbf{a}_{\mathrm{pa}(Y)\cap\mathbf{A}}, \{W(\mathbf{a}) \mid W \in \mathrm{pa}(Y) \setminus \mathbf{A}\})$$



$$, Y(a) = Y(a, M(a, W), W).$$

- Path-specific Potential outcome（16.2.1）
  - the treatment is set to one value a, with the path of interest. another value a' to other path.

$$Y(\pi, a, a') \equiv a \text{ if } Y = A$$
$$Y(\pi, a, a') \equiv Y(\{W(\pi, a, a') \mid W \in \mathrm{pa}^{\pi}(Y)\}, \{W(a') \mid W \in \mathrm{pa}^{\overline{\pi}}(Y)\})$$



$$\mathbb{E}[Y(a)] - \mathbb{E}[Y(\{A \to Z \to Y\}, a, a')] =$$
$$\mathbb{E}[Y(a)] - \mathbb{E}[Y(a', Z(a, M(a', W)), M(a', W), W)].$$

- g-formula/truncated factorization(16.2.4, 16.2.5)

$$p(\mathbf{V} \setminus \mathbf{A} \mid \mathrm{do}(\mathbf{a})) = \prod_{V \in \mathbf{V} \setminus \mathbf{A}} p(V \mid \mathrm{pa}(V))|_{\mathbf{A}=\mathbf{a}}$$

$$p(\{Y(\pi, \mathbf{a}, \mathbf{a}') | Y \in \mathbf{Y}\}) = \sum_{\mathbf{V} \setminus (\mathbf{A} \cup \mathbf{Y})} \prod_{V \in \mathbf{V} \setminus \mathbf{A}} p(V | \mathbf{a}_{\mathrm{pa}^{\pi}(V) \cap \mathbf{A}}, \mathbf{a}'_{\mathrm{pa}^{\overline{\pi}}(V) \cap \mathbf{A}}, \mathrm{pa}(V) \setminus \mathbf{A}).$$

Preliminaries & Terminologies

因果效应推断、估计量设计		
model misspecification bias	模型偏差	
outcome regression（OR）		Ma(2019)
propensity score （PS）	倾向得分	
nuisance function	“讨厌函数”、“憎函数”	
double robustness	双-鲁棒性	Ma(2019)
multiply robustness	多-鲁棒性	Eric(2012)

- ——A Robust and Efficient Approach to Causal Inference Based on Sparse Sufficient Dimension Reduction （Ma 2019）

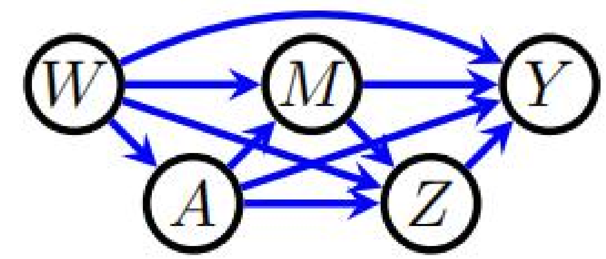
- ——Semi-parametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis（Eric 2012）

Preliminaries & Terminologies

估计量设计、估计量计算		
semi-parametric	半参数	Eric(2012)
efficient influence function	有效影响函数	Hahn(1998)、Eric(2012)
bootstrap	（有放回的采样）	
cross-fitting	（样本切分策略）	

- ——Semi-parametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis（Eric 2012）
- ——On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects （Hahn 1998）

Double Robustness



- Estimate：1.β_i=*E(Y(i))* (即，*A=i*)。2.β=（β_1,...,β_n）。*3.estimate* β

$$\beta^* := \arg\min_{\beta} \sum_{d=0}^{J} \mathbb{E}\left[\frac{I(D_i = d)}{\pi_d(X_i)} L\left(Y_i - \beta_d\right)\right]$$

- Propensity score：*f(A=a|W=w)*

Propensity score (PS): $\pi_d(x) := P(D_i = d | X_i = x)$

$L(v) = v^2$ $\quad \widehat{\beta}_d^{IPW} = \frac{1}{n}\sum_{i=1}^{n} \frac{I(D_i = d) Y_i}{\widehat{\pi}_d(X_i)}$

IPW:Inverse PS Weighting

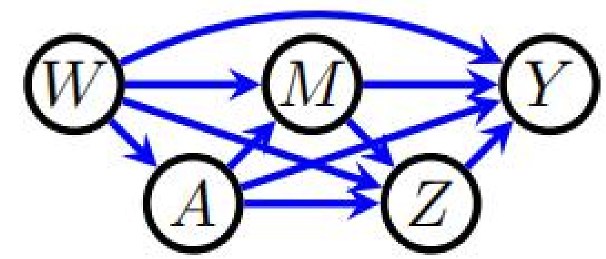
- Outcome regression:*E(Y|W,a,Z)*

$$\beta^* = \arg\min_{\beta} \sum_{d=0}^{J} \mathbb{E}\left[\mathbb{E}[L(Y - \beta_d) | X, D = d]\right]$$

OR function: $g_d(X) := \mathbb{E}[Y | X, D = d]$

$L(v) = v^2$ $\quad \beta_d^* = \mathbb{E}[Y(d)]$

Double Robustness



- *Efficient influence function:*

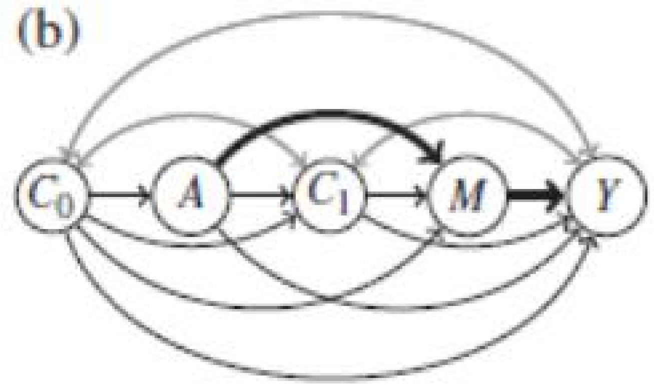
$$\eta_{i,d} = \frac{I\{D_i = d\}\{Y_i(d) - g_d(X_i)\}}{\pi_d(X_i)} + g_d(X_i)$$

- *Double robust Estimator:*

$$\widehat{\beta}_d^{DR} = n^{-1} \sum_{i=1}^{n} [\frac{I\{D_i = d\}\{Y_i(d) - \widehat{g}_d(X_i)\}}{\widehat{\pi}_d(X_i)} + \widehat{g}_d(X_i)]$$

- *Nuisance Function: to estimate PS function and OR function*

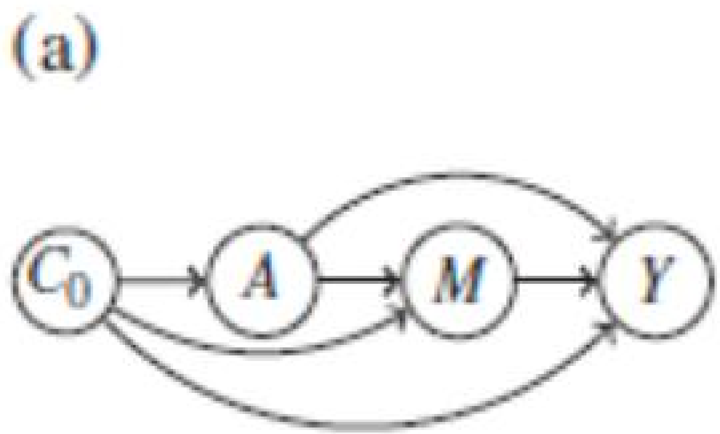
- This Paper
  - 因果图结构
    - 图节点：
      - C0：其他基础变量
      - A: indicator of Treatment
      - C1: 治疗的前6个月，药物毒性和结合力
      - M：治疗的第二个半年，药物结合力
      - Y: 病人的CD4计数水平
    - 假设：
      - Positivity：图中的每条路径的概率均大于0
      - Consistency
      - Ignorability：没有其他额外的混杂因子
  - 路径效应（A-M-Y）
    - P（AMY）：沿路径AMY的两个Po-out期望之差
  - 其中一个Po-out的半参数估计量
    - 构建估计量
      - 启发式构建
      - 统计泛函方式构建
    - 估计量的优点
      - Locally semi-para efficient：方差最小
      - Multiple robustness：尽可能减少Modeling Bias
      - Rate multiple robustness: 拓展了计算工具的选择范围
  - 数据验证
    - 模拟数据
    - PEPFAR（President's Emergency Plan for AIDS Relief）



Problem Description

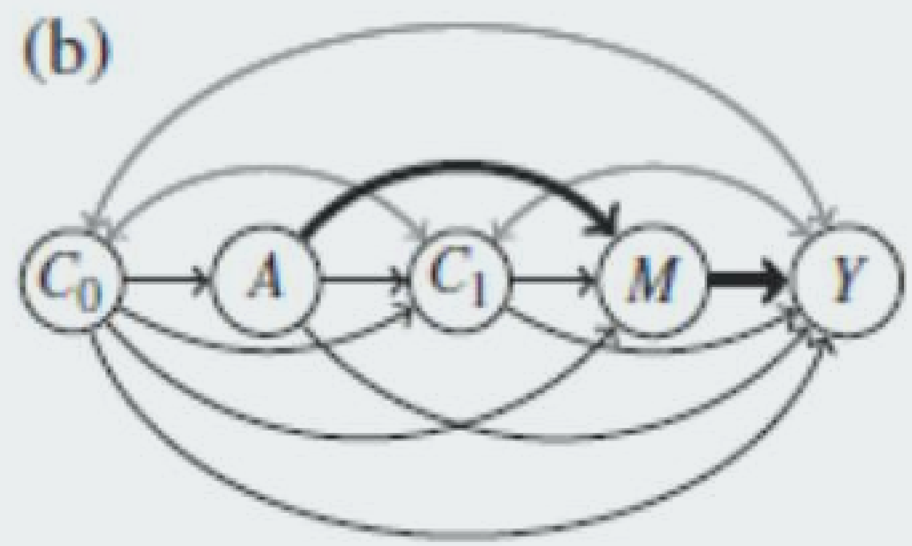
- 艾滋病疗法相关因素：1.药物；2.药物毒性（toxicity）；3.用药依从性（adherence）。
- 影响分析：1.药物毒性可能影响依从性；2.毒性以外的因素（如：meal restriction等），也可能影响依从性。
- **路径效应：纯粹由药物依从性（adherence），而非药物毒性（toxicity），带来的疗效。以病人的CD4计数水平为疗效指标。**
- 观察样本：Harvard PEPFAR（President's Emergency Plan for AIDS Relief），48345条观测样本。——注：有缺失数据问题，但不在本文讨论范围内。

Causal model & Path-specific effect

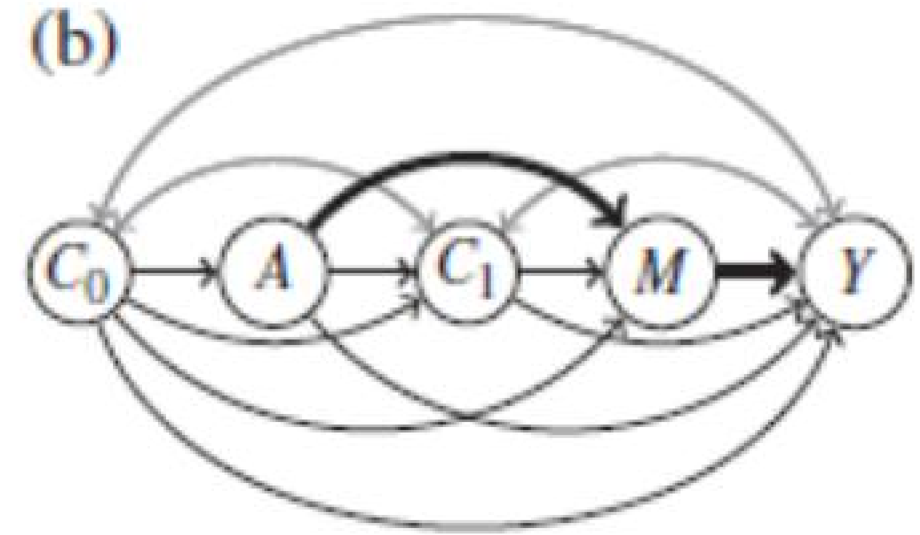


- 一个中介变量：M

- 多个中介变量：C1、M
- A对M、Y形成confounding
- C0、C1、Y之间有未知的confounding因素



Causal model & Path-specific effect



- C0：其他基础变量（随机向量）
- A: 不同逆转录疗法的indicator（0-1随机变量
    - a: comparison-level treatment 对照组
    - a‘：reference-level treatment 参考组
- C1: 在治疗的第一个半年，对药物毒性和用药依从度的观测数据。（随机向量）
- M：在治疗的第二个半年，仅对用药依从度水平的观测数据（随机变量）
- Y: 病人的CD4计数水平，表明HIV病毒的消除水平（随机变量）

Assumptions on Causal Graph

- Assumption1（Positivity、Overlap）：

$$0 < f_{A|C_0}(a \mid c_0) < 1 \qquad C_1^{\text{ratio}}(c_1, c_0) < \infty, \quad M^{\text{ratio}}(m, c_1, c_0) < \infty.$$

$$C_1^{\text{ratio}}(c_1, c_0) = f(c_1 \mid a', c_0)/f(c_1 \mid a, c_0)$$

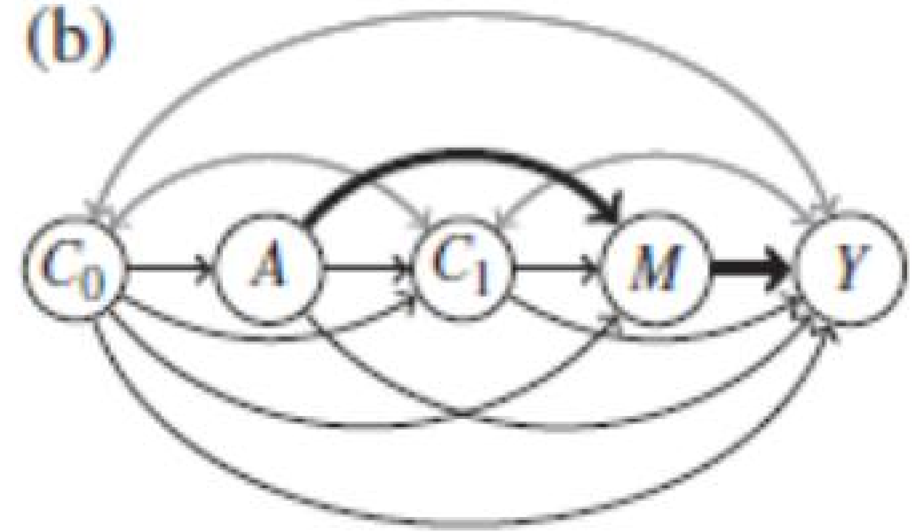
$$M^{\text{ratio}}(m, c_1, c_0) = f(m \mid c_1, a, c_0)/f(m \mid c_1, a', c_0).$$

- 0.每个病人都有一定的概率进入对照组或者参考组。

- 1. 对于对照组a的病人，前半年内，任何水平（Profile）的药物毒性和服药依从度都有一定概率出现。

- 2.对于参考组a’的病人，后半年内，任何水平（Profile）的药物依从度都有一定概率出现。

Assumptions on Causal Graph



- Assumption2（consistency）:

  if $W_1 = w_1$, then $W_2(w_1) = W_2$ almost everywhere.

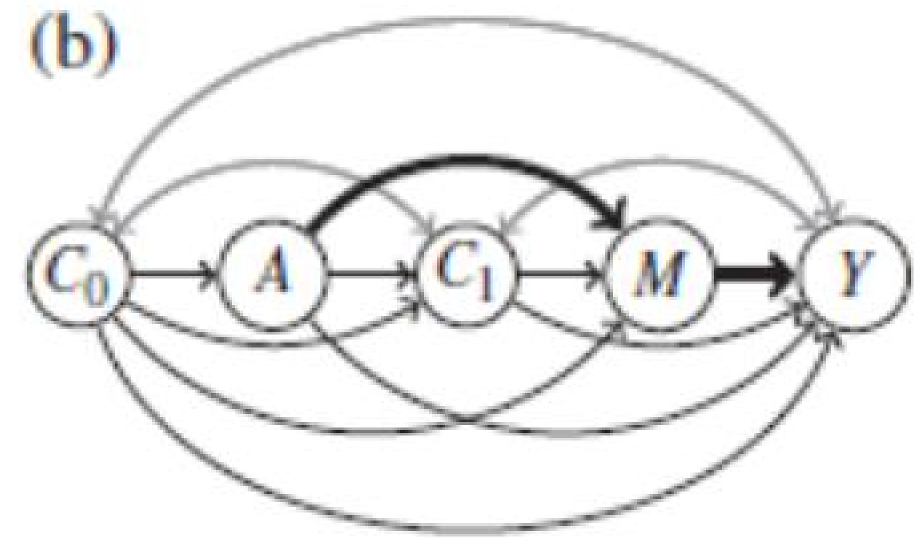
  $C_1(a^*)$, $M(a^*)$, $M(c_1, a^*)$ $\quad Y(m, c_1, a^*)$

- 含义：在W_1=w_1的情况下，Potential Outcome $W_2(w_1)$基本上就是观测值。

- 对于本文的数据来说，不管病人是“自然地（naturally）”分到对照组或是参考组，或者是被“干预”分配到某个组，没有区别。

- “这类假设是典型、且不可验证的”——本文作者

Assumptions on Causal Graph



- Assumption3（Ignorability、Unconfound）:

$$1.\ \{Y(m, a'), C_1(a')\} \perp A \mid C_0$$

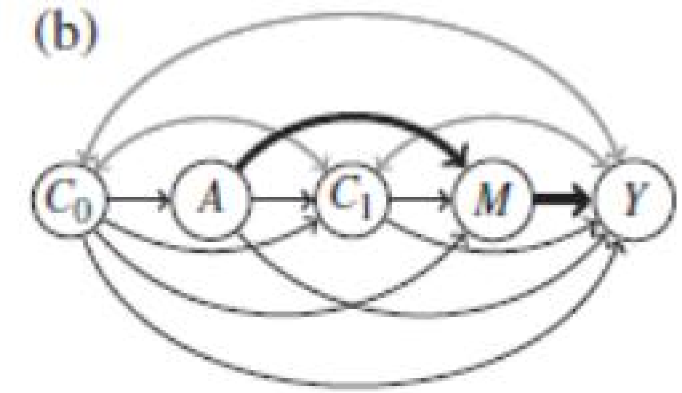
$$2.\ Y(m) \perp M \mid \{C_1, A, C_0\}$$

$$3.\ M(c_1, a) \perp \{C_1, A\} \mid C_0$$

$$4.\ \{Y(m, a'), C_1(a')\} \perp M(C_1, a) \mid C_0$$

- 1-3, 单世界条件独立；4，多世界条件独立。
- 注意：这里控制了A与C_1，A与M，A与Y，M与Y的unconfound，但没有假设C_0、C_1与Y之间必须是uncoufound。

Definition of path-specific effect



- 路径效应定义为两项Potential outcome之差（回顾16.2.1）

$$\mathcal{P}_{AMY} = E(Y[M\{C_1(a'), a\}, C_1(a'), a']) - E(Y[M\{C_1(a'), a'\}, C_1(a'), a']).$$

- 第一项是本文的核心待估参数，记为：

$$\beta_0 = E(Y[M\{C_1(a'), a\}, C_1(a'), a'])$$

- 又为（非参数形式）：

$$\beta_0 = \iiint_{c_0, c_1, m} E(Y \mid m, c_1, a', c_0)\, \mathrm{d}F(m \mid c_1, a, c_0)\, \mathrm{d}F(c_1 \mid a', c_0)\, \mathrm{d}F(c_0).$$

“半参数”，对上式中部分项的估计用到了参数模型（working model）

Semi-parametric estimators of β_0

- Estimator1：

$$\iiint_{m,c_1,c_0} \mathbb{E}(Y \mid m, c_1, e', c_0) dF_{M|C_1,E,C_0}(m \mid c_1, e, c_0) dF_{C_1|E,C_0}(c_1 \mid e', c_0) dF_{C_0}(c_0)$$

$$= \sum_{e^* \in \{e', e\}} \int_{y,m,c_1,c_0} y \frac{1_{e'}(e^*)}{f(e' \mid c_0)} \frac{f(m \mid c_1, e, c_0)}{f(m \mid c_1, e^*, c_0)} dF_{Y,M,C_1,E,C_0}(y, m, c_1, e^*, c_0)$$

$$= \mathbb{E}\left\{ \frac{1_{e'}(E)}{f(e' \mid C_0)} M^{ratio} Y \right\}.$$

$$\widehat{\beta}_d^{IPW} = \frac{1}{n} \sum_{i=1}^{n} \frac{I(D_i = d) Y_i}{\widehat{\pi}_d(X_i)}$$

$$\hat{\beta}_{\mathrm{a}} = \mathbb{P}_n \left\{ \frac{1_{a'}(A)}{\hat{f}(a' \mid C_0)} \hat{M}^{\text{ratio}}(M, C_1, C_0) Y \right\};$$

- 注：1. 增加$f(e'|c_0)$项，利用链式法则将条件分布按c_0、e'、c_1、m、Y的次序积分到Y。2.将待估项替换为估计项

Semi-parametric estimators of β_0

- Estimator2：

$$\iiint_{m,c_1,c_0} \mathbb{E}(Y \mid m, c_1, e', c_0) dF_{M|C_1,E,C_0}(m \mid c_1, e, c_0) dF_{C_1|E,C_0}(c_1 \mid e', c_0) dF_{C_0}(c_0)$$

$$= \sum_{e^* \in \{e', e\}} \int_{m,c_1,c_0} \mathbb{E}(Y \mid M, C_1, e', C_0) \frac{1_e(e^*)}{f(e^* \mid c_0)} \frac{f(c_1 \mid e', c_0)}{f(c_1 \mid e^*, c_0)} dF_{M,C_1,E,C_0}(m, c_1, e^*, c_0)$$

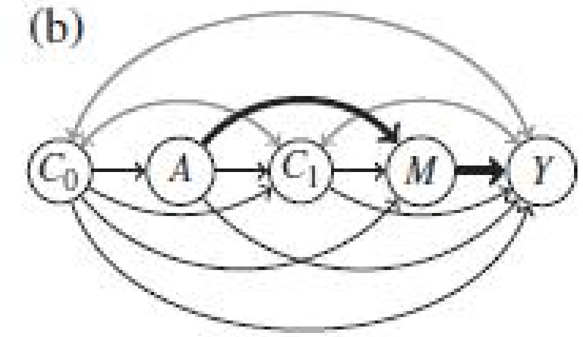
$$= \mathbb{E}\left[\frac{1_e(E)}{f(e \mid C_0)} (C_1^{\text{ratio}})^{-1} \mathbb{E}(Y \mid M, C_1, e', C_0)\right].$$

$$\widehat{\beta}_d^{IPW} = \frac{1}{n} \sum_{i=1}^{n} \frac{I(D_i = d) Y_i}{\widehat{\pi}_d(X_i)}$$

$$\hat{\beta}_{\mathrm{b}} = \mathbb{P}_n \left\{ \frac{1_a(A)}{\hat{f}(a \mid C_0)} \hat{C}_1^{\text{ratio}}(C_1, C_0) \hat{E}(Y \mid M, C_1, a', C_0) \right\}$$

- 注：1. 增加$f(e|c_0)$项，利用链式法则将条件分布按c_0、e、c_1、M的次序积分条件分布积分到M。2.将待估项替换为估计项。

- Estimate β_0 ,使用非参数的统计模型，M_{np}。



- 优点：flexible；缺点：“curse of dimensionality”。
- Estimate β_0 ,使用参数模型，parametric working model。

$$
\begin{aligned}
\hat{E}(Y \mid m, c_1, a^*, c_0) &= E^{\mathrm{W}}(Y \mid m, c_1, a^*, c_0; \hat{\gamma}_1) \\
\hat{f}_{A|C_0}(a \mid c_0) &= f^{\mathrm{W}}_{A|C_0}(a \mid c_0; \hat{\gamma}_4) \\
\hat{C}_1^{\text{ratio}}(c_1, c_0) &= C_1^{\text{ratio},W}(c_1, c_0; \hat{\gamma}_5) \\
\hat{M}^{\text{ratio}}(m, c_1, c_0) &= M^{\text{ratio},W}(m, c_1, c_0; \hat{\gamma}_6)
\end{aligned}
$$

- C_1、M连续时，可以使用贝叶斯公式估计：

$$
C_1^{\text{ratio}}(c_1, c_0) = \frac{f(c_1 \mid a', c_0)}{f(c_1 \mid a, c_0)} = \frac{f(a' \mid c_1, c_0)}{f(a \mid c_1, c_0)} \times \frac{f(a \mid c_0)}{f(a' \mid c_0)},
$$

$$
M^{\text{ratio}}(m, c_1, c_0) = \frac{f(m \mid a, c_1, c_0)}{f(m \mid a', c_1, c_0)} = \frac{f(a \mid m, c_1, c_0)}{f(a' \mid m, c_1, c_0)} \times \frac{f(a' \mid c_1, c_0)}{f(a \mid c_1, c_0)}.
$$

Semi-parametric estimators

- Estimator3：

$$\widehat{\beta}_d^{DR} = n^{-1}\sum_{i=1}^{n}[\frac{I\{D_i = d\}\{Y_i(d) - \widehat{g}_d(X_i)\}}{\widehat{\pi}_d(X_i)} + \widehat{g}_d(X_i)]$$

$$\begin{aligned}\hat{\beta}_{\mathrm{mr}} = \mathbb{P}_n\Bigg[& \frac{1_{a'}(A)}{\hat{f}(a' \mid C_0)}\hat{M}^{\mathrm{ratio}}(M, C_1, C_0)\{Y - \hat{B}(M, C_1, a', C_0)\} \\ & + \frac{1_a(A)}{\hat{f}(a \mid C_0)}\hat{C}_1^{\mathrm{ratio}}(C_1, C_0)\{\hat{B}(M, C_1, a', C_0) - \hat{B}'(C_1, a', a, C_0)\} \\ & + \frac{1_{a'}(A)}{\hat{f}(a' \mid C_0)}\{\hat{B}'(C_1, a', a, C_0) - \hat{B}''(a', a, C_0)\} + \hat{B}''(a', a, C_0)\Bigg].\end{aligned}$$

- 这个估计量得自于β_0的“有效影响函数（Efficient influence function）”EIF(β_0)。参见（2012，E.J. Tchetgen Tchetgen, I. Shpitser），β_0被视作一个M泛函（M-functional）

- 对于所有M_{np}中的渐进线性估计量，影响函数与EIF(β_0)相同的估计量，具有最小的渐进方差。

Efficient influence function:

$$\mathrm{EIF}(\beta_0) = \frac{1_{a'}(A)}{f(a' \mid C_0)} M^{\mathrm{ratio}}(M, C_1, C_0)\left\{Y - B(M, C_1, a', C_0)\right\}$$
$$+ \frac{1_{a}(A)}{f(a \mid C_0)} C_1^{\mathrm{ratio}}(C_1, C_0)\left\{B(M, C_1, a', C_0) - B'(C_1, a', a, C_0)\right\}$$
$$+ \frac{1_{a'}(A)}{f(a' \mid C_0)}\left\{B'(C_1, a', a, C_0) - B''(a', a, C_0)\right\} + \left\{B''(a', a, C_0) - \beta_0\right\}$$

Outcome regression function: $\beta_d^* = \mathbb{E}[Y(d)]$

$$B(m, c_1, a', c_0) = E(Y \mid m, c_1, a', c_0)$$

$$B'(c_1, a', a, c_0) = E\{E(Y \mid M, c_1, a', c_0) \mid c_1, a, c_0\}$$

$$B''(a', a, c_0) = E[E\{E(Y \mid M, C_1, a', c_0) \mid C_1, a, c_0\} \mid a', c_0]$$

Estimate β_0 ,使用参数模型，parametric working model。

$$B^{\mathrm{W}}(\gamma_1) \quad \hat{E}(Y \mid m, c_1, a^*, c_0) = E^{\mathrm{W}}(Y \mid m, c_1, a^*, c_0; \hat{\gamma}_1)$$

$$\hat{f}_{A|C_0}(a \mid c_0) = f^{\mathrm{W}}_{A|C_0}(a \mid c_0; \hat{\gamma}_4)$$

$$\hat{C}_1^{\mathrm{ratio}}(c_1, c_0) = C_1^{\mathrm{ratio},W}(c_1, c_0; \hat{\gamma}_5)$$

$$\hat{M}^{\mathrm{ratio}}(m, c_1, c_0) = M^{\mathrm{ratio},W}(m, c_1, c_0; \hat{\gamma}_6)$$

$$B'^{\mathrm{W}}(\gamma_2 \mid \gamma_1) = E^{\mathrm{W}}\{B^{\mathrm{W}}(\gamma_1) \mid M, C_1, a', C_0; \gamma_2\}$$

$$B''^{\mathrm{W}}(\gamma_3 \mid \gamma_1, \gamma_2) = E\{B'^{\mathrm{W}}(\gamma_1, \gamma_2) \mid C_1, a, C_0; \gamma_3\}.$$

$$\begin{aligned}
\hat{\beta}_{\mathrm{mr}} = \mathbb{P}_n \Bigg[& \frac{1_{a'}(A)}{\hat{f}(a' \mid C_0)} \hat{M}^{\mathrm{ratio}}(M, C_1, C_0)\left\{Y - \hat{B}(M, C_1, a', C_0)\right\} \\
& + \frac{1_a(A)}{\hat{f}(a \mid C_0)} \hat{C}_1^{\mathrm{ratio}}(C_1, C_0)\left\{\hat{B}(M, C_1, a', C_0) - \hat{B}'(C_1, a', a, C_0)\right\} \\
& + \frac{1_{a'}(A)}{\hat{f}(a' \mid C_0)}\left\{\hat{B}'(C_1, a', a, C_0) - \hat{B}''(a', a, C_0)\right\} + \hat{B}''(a', a, C_0) \Bigg].
\end{aligned}$$

Multiple robustness of Estimator3

- （使用parametric model）估计 $\hat{\beta}_{\mathrm{mr}}$ 的待估项时，只要以下之一成立，那么 $\hat{\beta}_{\mathrm{mr}}$ 是相容且渐进正态分布的：

$$
\begin{aligned}
&\text{(a)}\ \ \{\theta_M, f_{A|C_0}\} \in \{\theta_M^{\mathrm{W}}(\gamma_2, \gamma_6 \mid \gamma_1), f_{A|C_0}^{\mathrm{W}}(\gamma_4)\};\\
&\text{(b)}\ \ \{B, \theta_{C_1}, f_{A|C_0}\} \in \{B^{\mathrm{W}}(\gamma_1), \theta_{C_1}^{\mathrm{W}}(\gamma_3, \gamma_5 \mid \gamma_1, \gamma_2), f_{A|C_0}^{\mathrm{W}}(\gamma_4)\};\\
&\text{(c)}\ \ \{B, \theta_{C_1}, \theta_M\} \in \{B^{\mathrm{W}}(\gamma_1), \theta_{C_1}^{\mathrm{W}}(\gamma_3, \gamma_5 \mid \gamma_1, \gamma_2), \theta_M^{\mathrm{W}}(\gamma_2, \gamma_6 \mid \gamma_1)\},
\end{aligned}
$$

$$
\theta_M = \{B', M^{\mathrm{ratio}}\} \qquad \theta_{C_1} = \{B'', C_1^{\mathrm{ratio}}\}
$$

- 含义：在a),b),c)三种情形之一下，模型形式是正确的，那么估计量 $\hat{\beta}_{\mathrm{mr}}$不会产生model bias。相当于提高了model robustness。

- 并且，此处的a)条件，对应Estimator1 β_a；b)条件，对应 Estimator2 β_b。另外c)条件对应第三个估计量 β_c

Estimate β_0 ,使用非参数估计（e.g cross fitting）

$$\hat{\eta} = \{\hat{f}_{A|C_0}, \hat{\theta}_{C_1}, \hat{\theta}_M, \hat{B}\}$$

$$\hat{\eta}_{\mathrm{a}} = \{\hat{f}_{A|C_0}, \hat{\theta}_M\}$$

$$\hat{\eta}_{\mathrm{b}} = \{\hat{f}_{A|C_0}, \hat{\theta}_{C_1}, \hat{B}\}$$

$$\hat{\eta}_{\mathrm{c}} = \{\hat{\theta}_{C_1}, \hat{\theta}_M, \hat{B}\}$$

$$\theta_M = \{B', M^{\mathrm{ratio}}\} \qquad \theta_{C_1} = \{B'', C_1^{\mathrm{ratio}}\}$$

$$\begin{aligned}\hat{\beta}_{\mathrm{mr}} = \mathbb{P}_n\Bigg[& \frac{1_{a'}(A)}{\hat{f}(a' \mid C_0)} \hat{M}^{\mathrm{ratio}}(M, C_1, C_0)\{Y - \hat{B}(M, C_1, a', C_0)\} \\ & + \frac{1_a(A)}{\hat{f}(a \mid C_0)} \hat{C}_1^{\mathrm{ratio}}(C_1, C_0)\{\hat{B}(M, C_1, a', C_0) - \hat{B}'(C_1, a', a, C_0)\} \\ & + \frac{1_{a'}(A)}{\hat{f}(a' \mid C_0)}\{\hat{B}'(C_1, a', a, C_0) - \hat{B}''(a', a, C_0)\} + \hat{B}''(a', a, C_0)\Bigg].\end{aligned}$$

Rate multiple robustness of Estimator3

- （使用非参数方法“cross-fitting”估计 $\hat{\beta}_{\mathrm{mr}}$ 的待估项时），只要以下项的收敛速度满足以下条件，那么 $\hat{\beta}_{\mathrm{mr}}$ 是相且渐进正态分布的：

$$r_n(\hat{\eta}_x) r_n(\hat{\eta}_x^{\mathrm{c}}) = o(n^{1/2})$$

$$\hat{\eta}_x^{\mathrm{c}} = \hat{\eta} \setminus \hat{\eta}_x \text{ for } x \in \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$$

$r_n(\cdot)$ 该组估计量中最慢的收敛速度

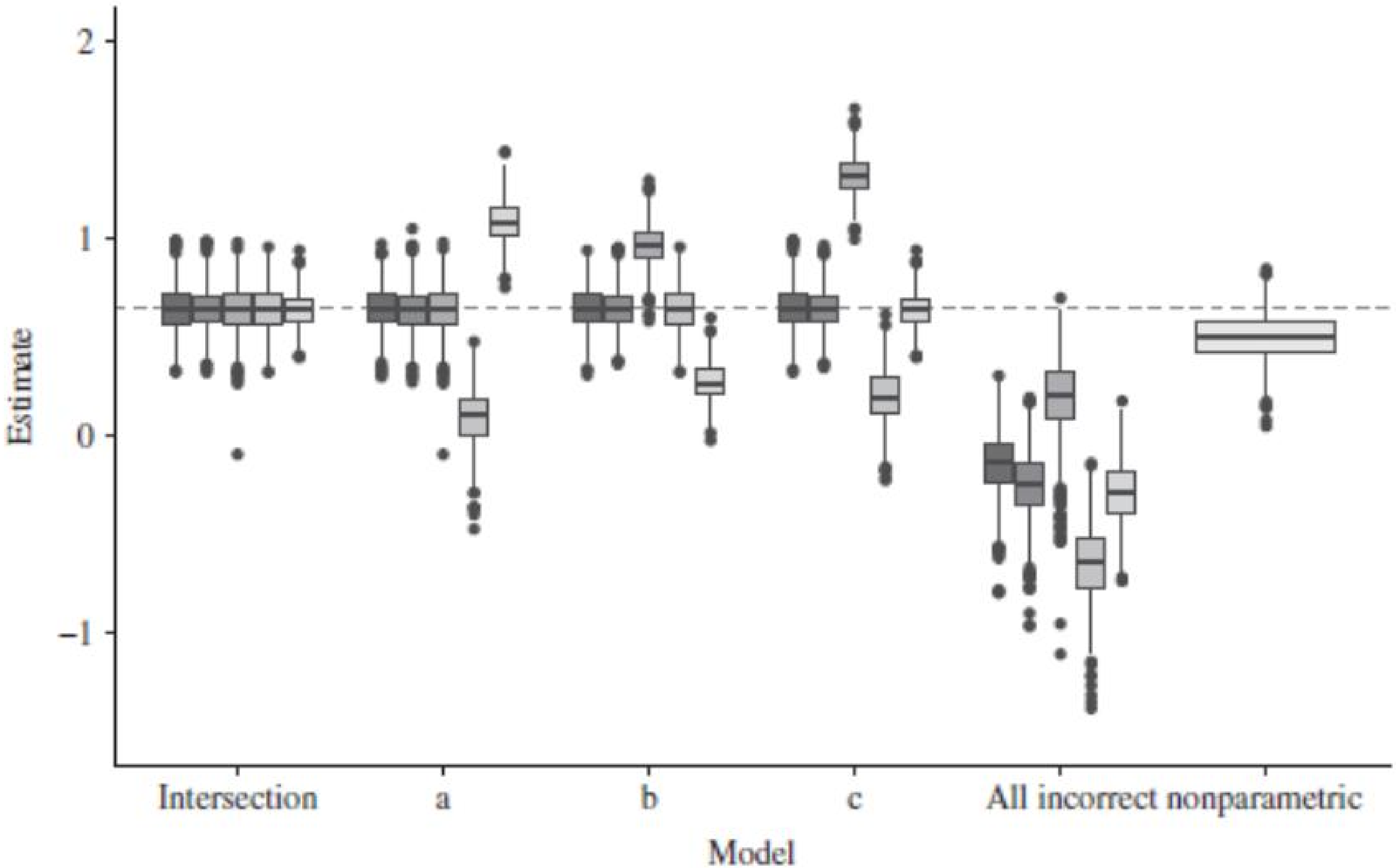
$$\hat{\eta}_{\mathrm{a}} = \{\hat{f}_{A|C_0}, \hat{\theta}_M\}$$

$$\hat{\eta}_{\mathrm{b}} = \{\hat{f}_{A|C_0}, \hat{\theta}_{C_1}, \hat{B}\}$$

$$\hat{\eta}_{\mathrm{c}} = \{\hat{\theta}_{C_1}, \hat{\theta}_M, \hat{B}\}$$

- 以上三族待估项中，只要有一族收敛速度足够快，Estimator3 $\hat{\beta}_{\mathrm{mr}}$ 是有效的估计量。

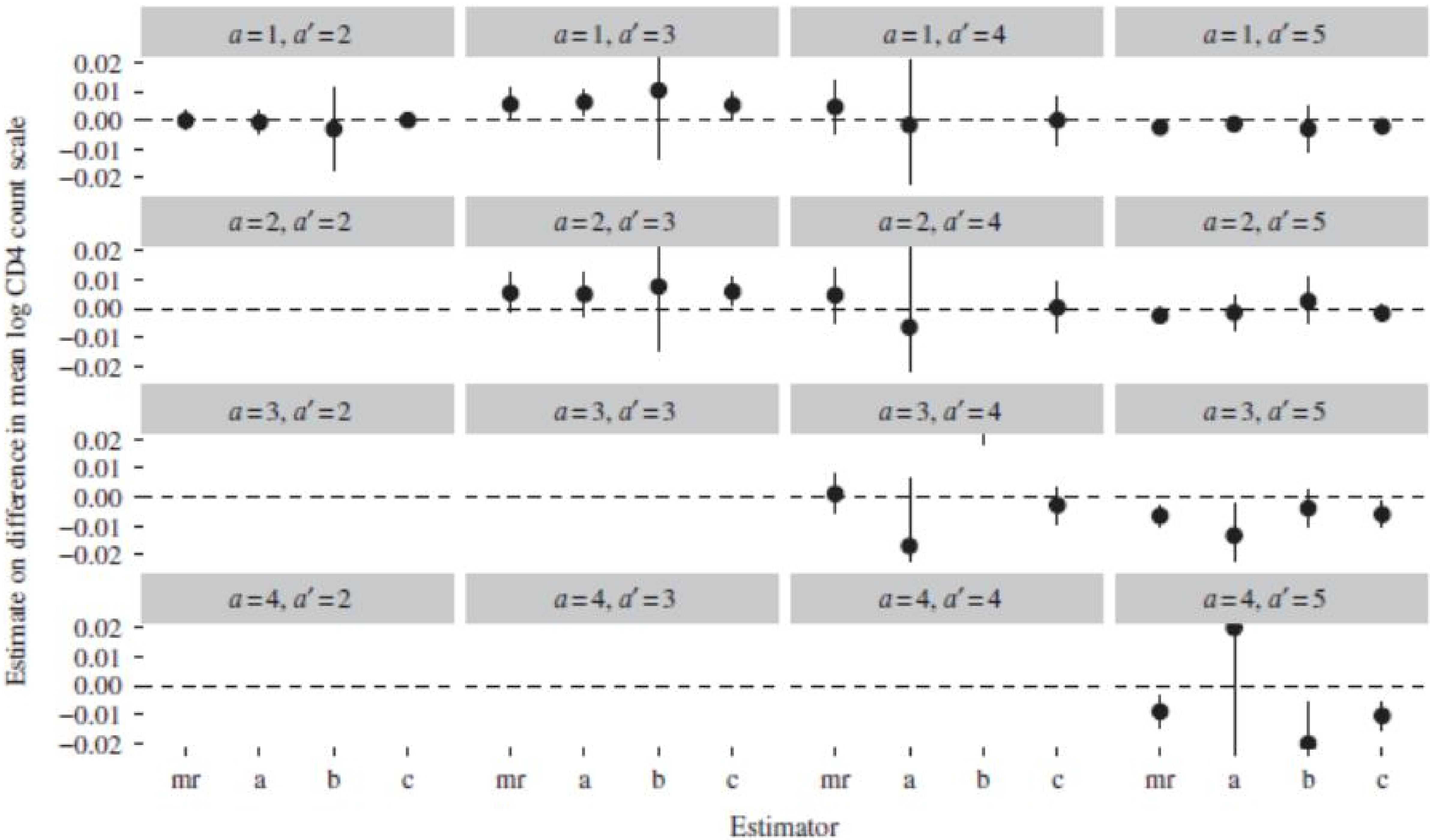
Simulation & PEPFAR results



Simulation & PEPFAR results

- 利用10000个样本容量的1000次人工数据实验，表明：Estimator3 β_{mr}（参数方法）对β_0的估计较为稳健；
- Estimator1、2（β_a、β_b）以及β_c对β_0的估计，容易出现model bias。
- Estimator3（非参数方法）在估计中出现偏差。作者认为这是非参数方法的收敛速度不够快的缘故。
- 在没有model bias情况下，β_a、β_b、β_c、β_{mr}的置信区间覆盖率在95%左右。
- β_{mr}（非参数方法）的置信区间覆盖率，当n=10000时，71%；当小样本时，接近95%

Simulation & PEPFAR results



使用估计量β_a、β_b、β_c、β_{mr}（参数方法）估计

Simulation & PEPFAR results

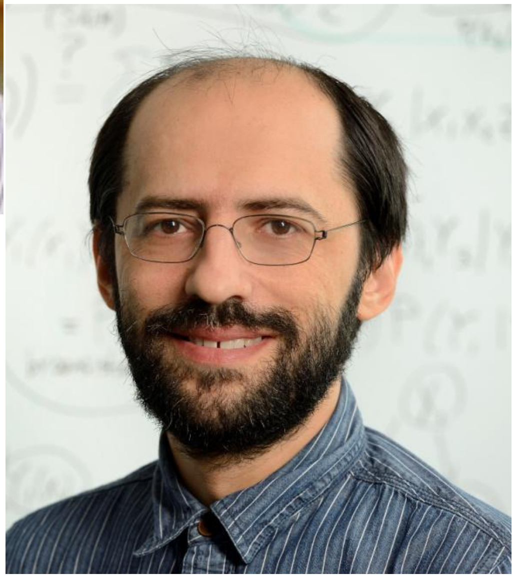
Table 1. *Estimated percentage of total effect on log CD4 count due to $\mathcal{P}_{AMY}$-specific effect*

	Baseline treatment			
Comparison treatment	2	3	4	5
1	−2	44*	7	−3*
2		−103	9	−4
3			4	−11*
4				−54*

*Significant path-specific effect ($\alpha = 0.05$). The treatments are coded as follows: 1 = AZT + 3TC + NVP, 2 = TDF + 3TC/FTC + EFV, 3 = AZT + 3TC + EFV, 4 = d4T + 3TC + NVP and 5 = TDF + 3TC/FTC + NVP, where 3TC = lamivudine, AZT = zidovudine, d4T = stavudine, EFV = efavirenz, FTC = emtricitabine, NVP = nevirapine and TDF = tenofovir.

About Authors & this paper

- Caleb Miles



- Ilyas Shpitser

- Phyllis Kanki

- Seema. Meloni

- Eric. Tchetgen Tchetgen

Further readings

- **Semi-parametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis**
- (2012, Eric J. Tchetgen Tchetgen, Ilya Shpitser)
- Quantifying an Adherence Path-Specific Effect of Antiretroviral Therapy in the Nigeria PEPFAR Program
- (2014, C. Miles, I.Shpister, P. Kanki, S.meloni, E.J.TT)
- **On semi-parametric estimation of a path-specific effect in the presence of mediator-outcome confounding**
- (2019, C. Miles, I.Shpister, P. Kanki, S.meloni, E.J.TT)

Further readings

- Counterfactual Graphical Models for Longitudinal Mediation Analysis With Unobserved Confounding
- (2013, I.Shpister)
- Characterization of parameters with a mixed bias property
- （2019,Rotnitzky)

Thanks