Causal Inference with Interference and Noncompliance in Two-Stage Randomized Controlled Trials

Kosuke Imai[†] Zhichao Jiang[†] Anup Malani[‡]

[†]Harvard [‡]Chicago

The Department of Statistics Colloquium

Harvard University

October 1, 2018

Methodological Motivation: Two-stage RCTs

- Causal inference revolution over the last three decades
- The first half of this revolution $\rightsquigarrow$ no interference between units

- In social sciences, interference is the rule rather than the exception
- Significant methodological progress over the last decade
- Experimental solution: two-stage randomized controlled trials (Hudgens and Halloran, 2008)

- We consider interference, both from encouragement to treatment and from treatment to outcome, in the presence of noncompliance

Empirical Motivation: Indian Health Insurance Experiment

- What are the health and financial effects of expanding a national health insurance program?
- RSBY (Rashtriya Swasthya Bima Yojana) subsidizes health insurance for "below poverty line" (BPL) Indian households
  - Monthly household income below ₹900 (rural) / 1,100 (urban) in Karnakata
  - Pays for hospitalization expenses
  - No deductible or copay with the annual limit of ₹30,000
  - Household pays ₹30 for smart card fee
  - Government pays about ₹200 for insurance premium in Karnakata
- We conduct an RCT to evaluate the impact of expanding RSBY to non-poor (i.e., APL or above poverty line) households
- Does health insurance have spillover effects on non-beneficiaries?

Study Design

- Sample: 10,879 households in 435 villages
- Experimental conditions:
  - A. Opportunity to enroll in RSBY essentially for free
  - B. No intervention
- Time line:
  1. September 2013 – February 2014: Baseline survey
  2. April – May 2015: Enrollment
  3. September 2016 – January 2017: Endline survey
- Two stage randomization:

Mechanisms	Village prop.	Treatment	Control
High	50%	80%	20%
Low	50%	40%	60%

Causal Inference and Interference between Units

1. Causal inference without interference between units
   - Potential outcomes: $Y_i(1)$ and $Y_i(0)$
   - Observed outcome: $Y_i = Y_i(T_i)$
   - Causal effect: $Y_i(1) - Y_i(0)$

2. Causal inference with interference between units
   - Potential outcomes: $Y_i(t_1, t_2, \ldots, t_N)$
   - Observed outcome: $Y_i = Y_i(T_1, T_2, \ldots, T_N)$
   - Causal effects:
     - Direct effect $= Y_i(T_i = 1, \mathbf{T}_{-i} = \mathbf{t}) - Y_i(T_i = 0, \mathbf{T}_{-i} = \mathbf{t})$
     - Spillover effect $= Y_i(T_i = t, \mathbf{T}_{-i} = \mathbf{t}) - Y_i(T_i = t, \mathbf{T}_{-i} = \mathbf{t}')$

Fundamental problem of causal infernece
$\rightsquigarrow$ only one potential outcome is observed

Two-stage Randomized Experiments

- Individuals (households): $i = 1, 2, \ldots, N$
- Blocks (villages): $j = 1, 2, \ldots, J$
- Size of block j: n_j where $N = \sum_{j=1}^{J} n_j$

- Binary treatment assignment mechanism: $A_j \in \{0, 1\}$
- Binary encouragement to receive treatment: $Z_{ij} \in \{0, 1\}$
- Binary treatment indicator: $D_{ij} \in \{0, 1\}$
- Observed outcome: Y_{ij}

- Partial interference assumption: No interference across blocks
  - Potential treatment and outcome: $D_{ij}(\mathbf{z}_j)$ and $Y_{ij}(\mathbf{z}_j)$
  - Observed treatment and outcome: $D_{ij} = D_{ij}(\mathbf{Z}_j)$ and $Y_{ij} = Y_{ij}(\mathbf{Z}_j)$
- Number of potential values reduced from 2^N to 2^{n_j}

Intention-to-Treat Analysis: Causal Quantities of Interest

- Average outcome under the treatment $Z_{ij} = z$ and the assignment mechanism $A_j = a$:

$$\overline{Y}_{ij}(z, a) = \sum_{\mathbf{z}_{-i,j}} Y_{ij}(Z_{ij} = z, \mathbf{Z}_{-i,j} = \mathbf{z}_{-i,j}) \mathbb{P}_a(\mathbf{Z}_{-i,j} = \mathbf{z}_{-i,j} \mid Z_{ij} = z)$$

- Average direct effect of encouragement on outcome:

$$\text{ADE}^Y(a) = \frac{1}{N} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left\{ \overline{Y}_{ij}(1, a) - \overline{Y}_{ij}(0, a) \right\}$$

- Average spillover effect of encouragement on outcome:

$$\text{ASE}^Y(z) = \frac{1}{N} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \left\{ \overline{Y}_{ij}(z, 1) - \overline{Y}_{ij}(z, 0) \right\}$$

- Horvitz-Thompson estimator for unbiased estimation

Effect Decomposition

- Average total effect of encouragement on outcome:

$$\text{ATE}^Y = \frac{1}{N}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\left\{\overline{Y}_{ij}(1,1) - \overline{Y}_{ij}(0,0)\right\}$$

- Total effect = Direct effect + Spillover effect:

$$\text{ATE}^Y = \text{ADE}^Y(1) + \text{ASE}^Y(0) = \text{ADE}^Y(0) + \text{ASE}^Y(1)$$

- In a two-stage RCT, we have an unbiased estimator,

$$\mathbb{E}\left[\frac{\sum_{j=1}^{J}\mathbf{1}\{A_j = a\}\frac{n_j}{N}\frac{\sum_{i=1}^{n_j} Y_{ij}\mathbf{1}\{Z_{ij}=z\}}{\sum_{i=1}^{n_j}\mathbf{1}\{Z_{ij}=z\}}}{\frac{1}{J}\sum_{j=1}^{J}\mathbf{1}\{A_j = a\}}\right] = \frac{1}{N}\sum_{j=1}^{J}\sum_{i=1}^{n_j}\overline{Y}_{ij}(z,a)$$

- Halloran and Struchiner (1995), Sobel (2006), Hudgens and Halloran (2008)

Complier Average Direct Effect

- Goal: Estimate the treatment effect rather than the ITT effect
- Use randomized encouragement as an instrument
  1. Monotonicity: $D_{ij}(1, \mathbf{z}_{-i,j}) \geq D_{ij}(0, \mathbf{z}_{-i,j})$ for any $\mathbf{z}_{-i,j}$
  2. Exclusion restriction: $Y_{ij}(\mathbf{z}_j, \mathbf{d}_j) = Y_{ij}(\mathbf{z}'_j, \mathbf{d}_j)$ for any $\mathbf{z}_j$ and $\mathbf{z}'_j$

- Compliers: $C_{ij}(\mathbf{z}_{-i,j}) \ = \ \mathbf{1}\{D_{ij}(1, \mathbf{z}_{-i,j}) = 1, D_{ij}(0, \mathbf{z}_{-i,j}) = 0\}$
- Complier average direct effect of encouragement (CADE(z, a)):

$$\frac{\sum_{j=1}^{J} \sum_{i=1}^{n_j} \{Y_{ij}(1, \mathbf{z}_{-i,j}) - Y_{ij}(0, \mathbf{z}_{-i,j})\} C_{ij}(\mathbf{z}_{-i,j}) \mathbb{P}_a(\mathbf{Z}_{-i,j} = \mathbf{z}_{-i,j} \mid Z_{ij} = z)}{\sum_{j=1}^{J} \sum_{i=1}^{n_j} C_{ij}(\mathbf{z}_{-i,j}) \mathbb{P}_a(\mathbf{Z}_{-i,j} = \mathbf{z}_{-i,j} \mid Z_{ij} = z)}$$

- We propose a consistent estimator of the CADE

Key Identification Assumption

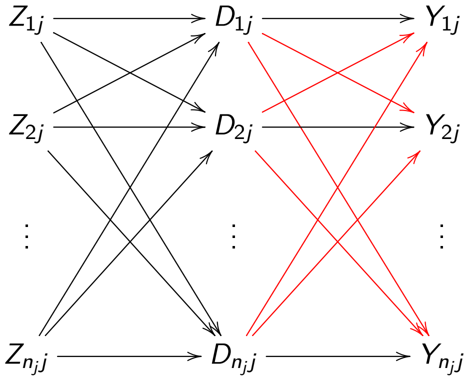
- Two causal mechanisms:
  - Z_{ij} affects Y_{ij} through D_{ij}
  - Z_{ij} affects Y_{ij} through $\mathbf{D}_{-i,j}$
- Idea: if Z_{ij} does not affect D_{ij}, it should not affect Y_{ij} through $\mathbf{D}_{-i,j}$

Assumption (Restricted Interference for Noncompliers)

If a unit has $D_{ij}(1, \mathbf{z}_{-i,j}) = D_{ij}(0, \mathbf{z}_{-i,j}) = d$ for any given $\mathbf{z}_{-i,j}$, it must also satisfy $Y_{ij}(d, \mathbf{D}_{-i,j}(Z_{ij} = 1, \mathbf{z}_{-i,j})) = Y_{ij}(d, \mathbf{D}_{-i,j}(Z_{ij} = 0, \mathbf{z}_{-i,j}))$

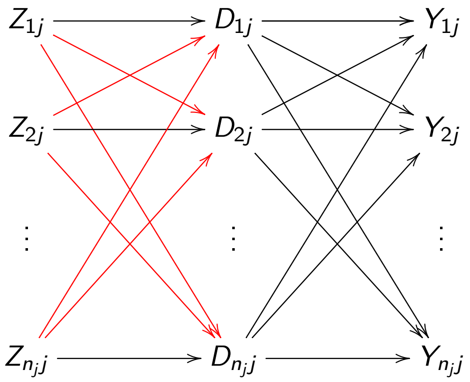
Scenario I: No Spillover Effect of the Treatment Receipt on the Outcome

$$Y_{ij}(d_{ij}, \mathbf{d}_{-i,j}) \;=\; Y_{ij}(d_{ij}, \mathbf{d}'_{-i,j})$$



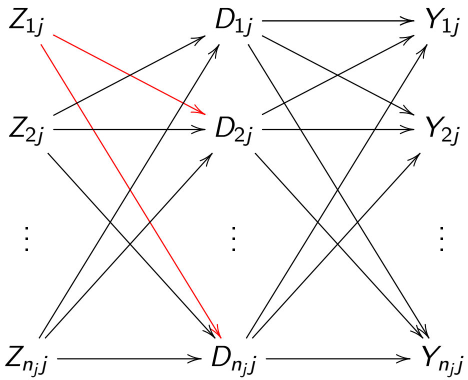
Scenario II: No Spillover Effect of the Treatment Assignment on the Treatment Receipt

$$D_{ij}(z_{ij}, \mathbf{z}_{-i,j}) \ = \ D_{ij}(z_{ij}, \mathbf{z}'_{-i,j})$$ (Kang and Imbens, 2016)



Scenario III: Limited Spillover Effect of the Treatment Assignment on the Treatment Receipt

$$\text{If } D_{ij}(1, \mathbf{z}_{-i,j}) = D_{ij}(0, \mathbf{z}_{-i,j}) \text{ for any given } \mathbf{z}_{-i,j},$$
$$\text{then } D_{i'j}(1, \mathbf{z}_{-i,j}) = D_{i'j}(0, \mathbf{z}_{-i,j}) \text{ for all } i' \neq i$$



Identification and Consistent Estimation

1. Identification: monotonicity, exclusion restriction, restricted interference for noncompliers

$$\lim_{n_j \to \infty} \mathrm{CADE}(z, a) \; = \; \lim_{n_j \to \infty} \frac{\mathrm{ADE}^Y(a)}{\mathrm{ADE}^D(a)}$$

2. Consistent estimation: additional restriction on interference (e.g., Savje et al.)

$$\frac{\widehat{\mathrm{ADE}}^Y(a)}{\widehat{\mathrm{ADE}}^D(a)} \xrightarrow{p} \lim_{n_j \to \infty, J \to \infty} \mathrm{CADE}(z, a)$$

Randomization Inference

- Variance is difficult to characterize

Assumption (Stratified Interference (Hudgens and Halloran. 2008))

$Y_{ij}(z_{ij}, \mathbf{z}_{-i,j}) = Y_{ij}(z_{ij}, \mathbf{z}'_{-i,j})$ and $D_{ij}(z_{ij}, \mathbf{z}_{-i,j}) = D_{ij}(z_{ij}, \mathbf{z}'_{-i,j})$ *if* $\sum_{i'=1}^{n_j} z_{ij} = \sum_{i=1}^{n_j} z'_{ij}$

- Under stratified interference, our estimand simplifies to,

$$\begin{aligned} & \mathsf{CADE}(a) \\ = \quad & \frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\{Y_{ij}(1,a) - Y_{ij}(0,a)\}\mathbf{1}\{D_{ij}(1,a)=1, D_{ij}(0,a)=0\}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\mathbf{1}\{D_{ij}(1,a)=1, D_{ij}(0,a)=0\}} \end{aligned}$$

- Compliers: $C_{ij} = \mathbf{1}\{D_{ij}(1,a)=1, D_{ij}(0,a)=0\}$
- Consistent estimation possible without additional restriction
- We propose an approximate asymptotic variance estimator

Connection to the Two-stage Least Squares Estimator

- The model:

$$
\begin{aligned}
Y_{ij} &= \sum_{a=0}^{1} \alpha_a \mathbf{1}\{A_j = a\} + \sum_{a=0}^{1} \underbrace{\beta_a}_{\text{CADE}} D_{ij} \mathbf{1}\{A_j = a\} + \epsilon_{ij} \\
D_{ij} &= \sum_{a=0}^{1} \gamma_a \mathbf{1}\{A_j = a\} + \sum_{a=0}^{1} \delta_a Z_{ij} \mathbf{1}\{A_j = a\} + \eta_{ij}
\end{aligned}
$$

- Weighted two-stage least squares estimator:

$$
w_{ij} = \frac{1}{\Pr(A_j)\Pr(Z_{ij} \mid A_j)}
$$

- Transforming the outcome and treatment: multiplying them by $n_j J / N$
- Randomization-based variance is equal to the weighted average of cluster-robust HC2 and individual-robust HC2 variances

Complier Average Spillover Effect

- Under stratified interference, we can define the average spillover effect for compliers

Assumption (Monotonicity with respect to Assignment Mechanism)

$$D_{ij}(z, 1) \;\geq\; D_{ij}(z, 0)$$

- Compliers: $\mathbf{1}\{D_{ij}(z,1) = 1, D_{ij}(z,0) = 0\}$
- Complier Average Spillover Effect (CASE):

$$\begin{aligned}&\text{CASE}(z)\\ = \;\; & \frac{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\{Y_{ij}(z,1) - Y_{ij}(z,0)\}\mathbf{1}\{D_{ij}(z,1) = 1, D_{ij}(z,0) = 0\}}{\sum_{j=1}^{J}\sum_{i=1}^{n_j}\mathbf{1}\{D_{ij}(z,1) = 1, D_{ij}(z,0) = 0\}}.\end{aligned}$$

- Consistent estimation:

$$\frac{\widehat{\text{ASE}}^{Y}(z)}{\widehat{\text{ASE}}^{D}(z)} \xrightarrow{p} \lim_{n_j \to \infty, J \to \infty} \text{CASE}(z)$$

Simulation Setup

- Two assignment mechanisms ($A_j = 0$: 40%, $A_j = 1$: 60%):
  1. $\Pr(Z_{ij} = 1 \mid A_j = 0) = 0.4$
  2. $\Pr(Z_{ij} = 1 \mid A_j = 1) = 0.6$

- Compliance status:

$$C_{ij}(a) = \begin{cases} \text{complier} & \text{if } D_{ij}(1,a) = 1, D_{ij}(0,a) = 0 \\ \text{always} - \text{taker} & \text{if } D_{ij}(1,a) = D_{ij}(0,a) = 1 \\ \text{never} - \text{taker} & \text{if } D_{ij}(1,a) = D_{ij}(0,a) = 0 \end{cases}$$

- Spillover effect of encouragement on treatment $\rightsquigarrow$ complier status proportions (complier, always-taker, never-taker)
  1. $a = 0$: (40%, 30%, 30%)
  2. $a = 1$: (60%, 20%, 20%)
- No spillover effect: $C_{ij}(1) = C_{ij}(0)$ for all i, j and (50%, 30%, 20%)

- No spillover effect of treatment on outcome

$$Y_{ij}(d_{ij}=0) \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$
$$Y_{ij}(1) - Y_{ij}(0) \overset{\text{indep.}}{\sim} \mathcal{N}(\theta_j, \sigma^2)$$

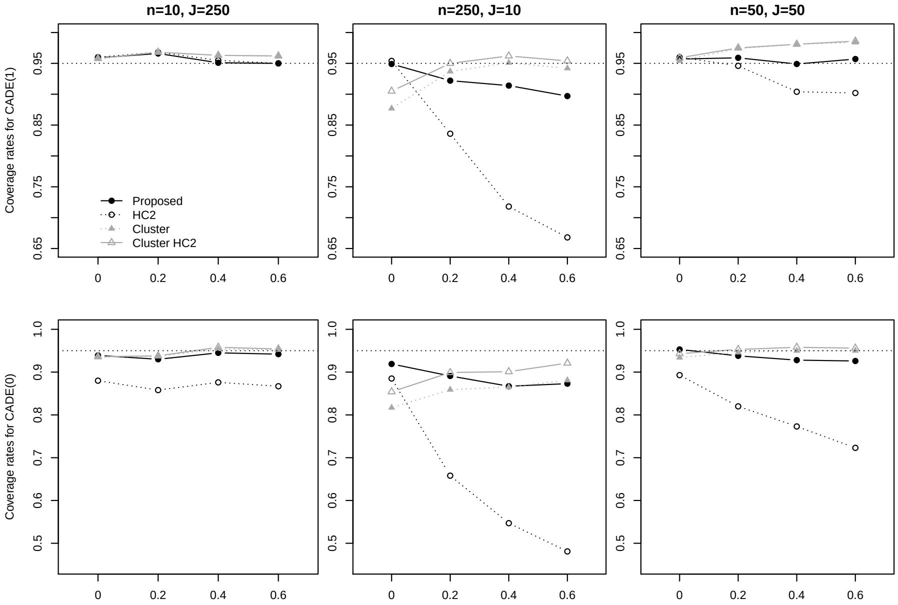
- Spillover effect of treatment on outcome: $\rightsquigarrow$ stratified interference

$$Y_{ij}(0, \mathbf{d}_{-i,j}) \overset{\text{indep.}}{\sim} \mathcal{N}\left(\frac{\beta}{n_j}\sum_{i'} d_{i'j},\ 1\right)$$
$$Y_{ij}(1, \mathbf{d}_{-i,j}) - Y_{ij}(0, \mathbf{d}_{-i,j}) \overset{\text{indep.}}{\sim} \mathcal{N}(\theta_j, \sigma^2)$$

- $\theta_j \overset{\text{indep.}}{\sim} \mathcal{N}(\theta, \omega^2)$

- Vary intracluster correlation coefficient $\rho = \omega^2/(\sigma^2+\omega^2)$
- Vary cluster size n and number of clusters J

Results: Both Spillover Effects Present



Results: Indian Health Insurance Experiment

- A household is more likely to enroll in RSBY if a large number of households are given the opportunity

Average Spillover Effects	Treatment	Control
Individual-weighted	0.086 (s.e. = 0.053)	0.045 (s.e. = 0.028)
Block-weighted	0.044 (s.e. = 0.018)	0.031 (s.e. = 0.021)

- Households will have greater hospitalization expenditure if few households are given the opportunity

Complier Average Direct Effects	High	Low
Individual-weighted	−1649 (s.e. = 1061)	1984 (s.e. = 1215)
Block-weighted	−485 (s.e. = 1258)	3752 (s.e. = 1652)

Concluding Remarks

- In social science research,
  1. people interact with each other $\rightsquigarrow$ interference
  2. people don't follow instructions $\rightsquigarrow$ noncompliance

- Two-stage randomized controlled trials:
  1. randomize assignment mechanisms across clusters
  2. randomize treatment assignment within each cluster

- Our contributions:
  1. Identification condition for complier average direct effects
  2. Consistent estimator for CADE and its variance
  3. Connections to regression and instrumental variables
  4. Application to the India health insurance experiment
  5. Implementation as part of R package experiment

Send comments and suggestions to
Imai@Harvard.Edu