Estimating average causal effects under general interference between units

Peter M. Aronow and Cyrus Samii

Yale University and New York University

March 2, 2012

- Randomized experiments often involve treatments that may induce "interference between units"
- Interference: the outcome for unit i depends on the treatment assigned to unit j. If we administer a treatment to unit j, what are the effects on unit i?
- Traditionally a nuisance, but now a topic of study – in the study of spillovers, equilibrium adjustment, networks, etc.
- Recent work in non-parametric inference focuses on hypothesis testing or estimation in hierarchical (i.e., multilevel) interference settings. We develop a theory of design-based estimation under general interference.

What's out there?

Figure 2: Section of Village with geographical clusters

Notes: The solid white lines delimit a geographical cluster. A square represents the location of a T_1 household, a star represents a T_2 household and a dot represents a control household in a control cluster. A triangle represents a control household in a treated cluster (either T_1 or T_2).

(Giné & Mansuri, 2011)

$$Y_{ijt} = a + \beta_1 \cdot T_{1it} + \beta_2 \cdot T_{2it} + X'_{ijt}\delta + \sum_d (\gamma_d \cdot N^T_{dit}) + \sum_d (\phi_d \cdot N_{dit}) + u_i + e_{ijt}.$$

school i in year t of the program.[26] Given the total number of children attending primary school within a certain distance from the school, the number of these attending schools assigned to treatment is exogenous and random. Since any independent effect of local school density is captured in the N_{dit} terms, the γ_d coefficients measure the deworming treatment externalities across schools.

(Miguel & Kremer, 2004, 175-6)

- Linear approximation of indirect exposure from to N^T_{di}.
- Requires extrapolation, since $Pr(N^T_{di} = n) = 0$ for some i, n.
- Even under generous assumptions, fixed effects would not aggregate to ATE (Angrist & Pischke, 2009).
- Subtle ratio estimation biases for finite samples.
- Variance estimation? Not clear ex ante, given complex dependencies between units.

- We provide a nonparametric *design-based* method for estimating average causal effects, including (but not limited to):
- Direct effect of assigning a unit to treatment
- Indirect effects of, e.g., a unit's peer being assigned to treatment
- More complex effects (e.g., effect of having a majority of proximal peers treated)
- The researcher must have knowledge of two characteristics:
  - The design of the experiment. What is the probability profile over all possible treatment assignments?
  - The exposure model. How do treatment assignments map onto actual exposures, direct or indirect?
- Methods are based on Horvitz-Thompson (HT) estimation (sample theoretic).

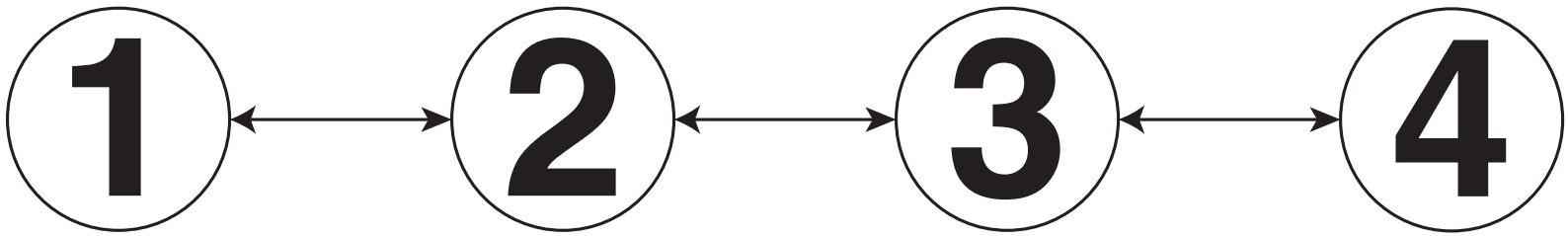
Method summary:

- The analyst specifies an exposure model, converting vectors of assigned treatments to vectors of actual exposures
- The analyst computes the *exact* probabilities that each unit will receive a given exposure
- The probabilities yield a simple, unbiased estimator of average causal effects

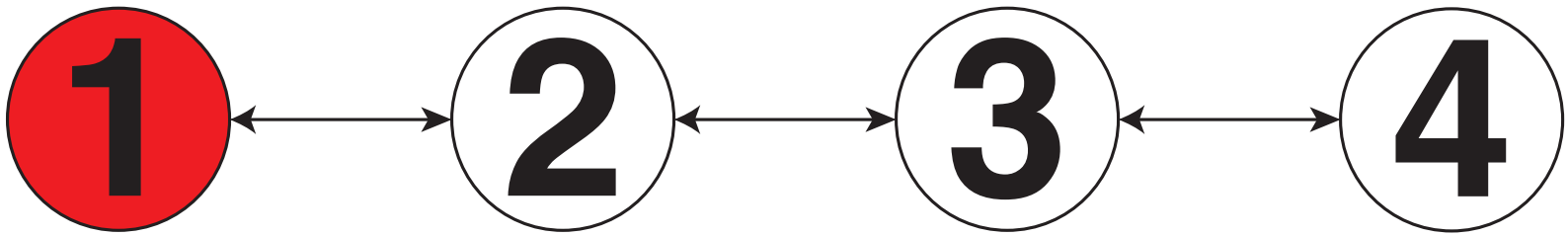
What you should remember from this presentation, if nothing else:

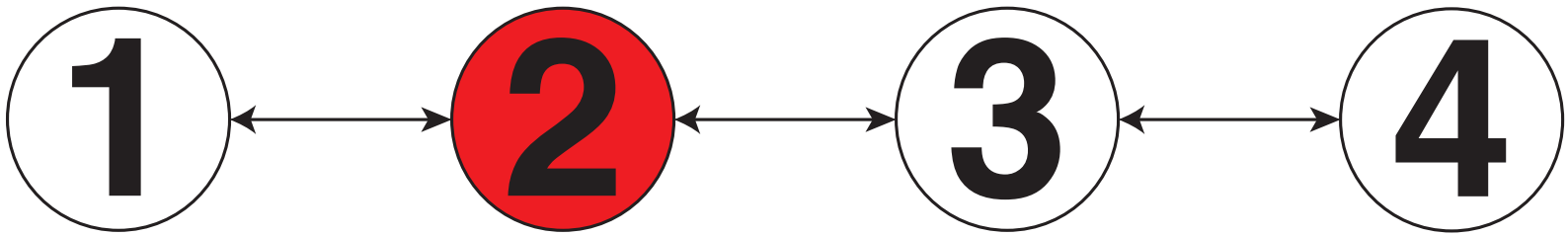
- Equal probability randomization does **NOT** imply equal probability of exposure
- Common naive methods ignoring these unequal probabilities (e.g., difference-in-means, regression) can lead to bias, even asymptotically

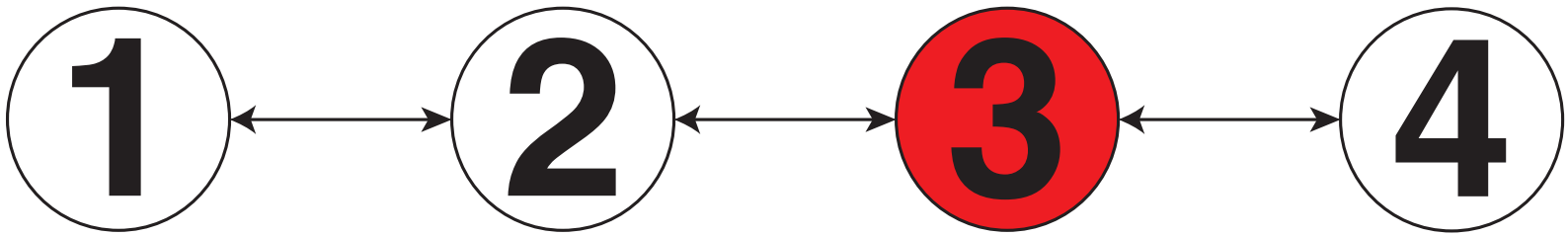
- To ground concepts, we provide a simple running example
- Consider a randomized experiment performed on a finite population of four units in a simple, fixed network:

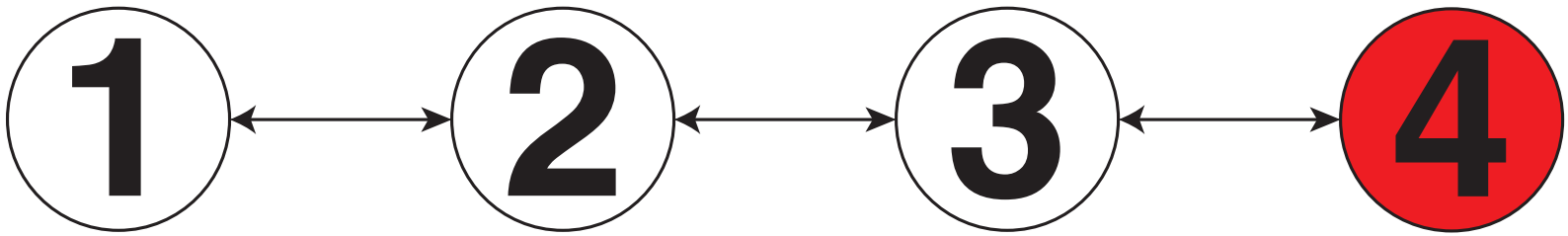

1
2
3
4

- One of these units is assigned to receive an campaign advertisement and the other three are assigned to control, equal probability
- We want to estimate the effects of advertising on opinion
- There are four possible randomizations $\mathbf{z}$:


1
2
3
4


1
2
3
4


1
2
3
4

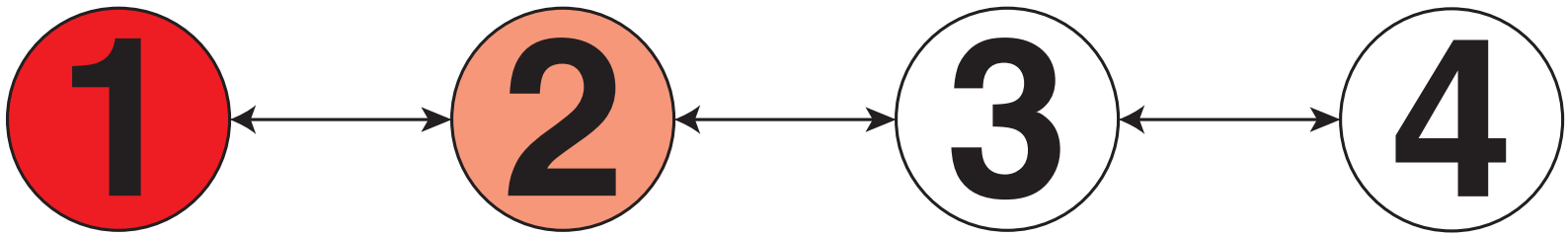

1
2
3
4

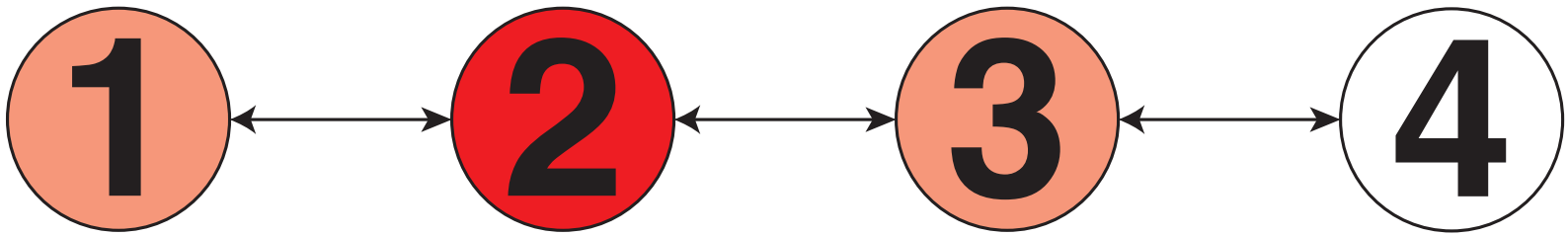
- So we have exact knowledge of the randomization scheme.
- But what of the exposure model? This requires researcher discretion. How do we model exposure to a treatment?
- One example.

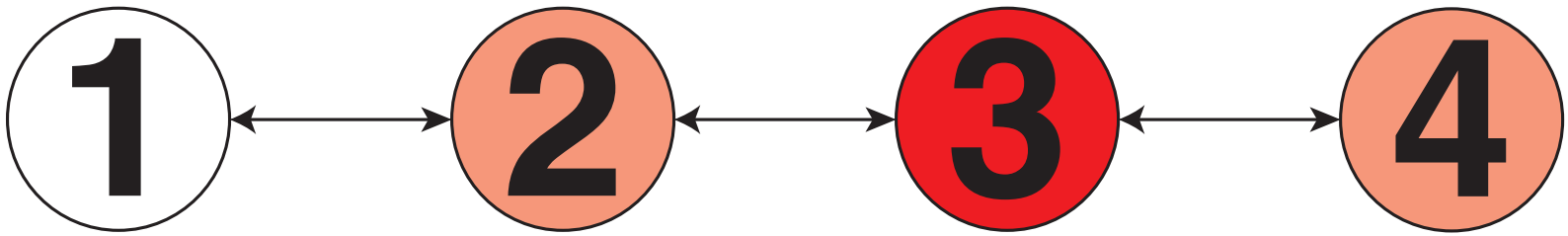
- Direct exposure means that you have been treated.
- Indirect exposure means that a peer has been treated.

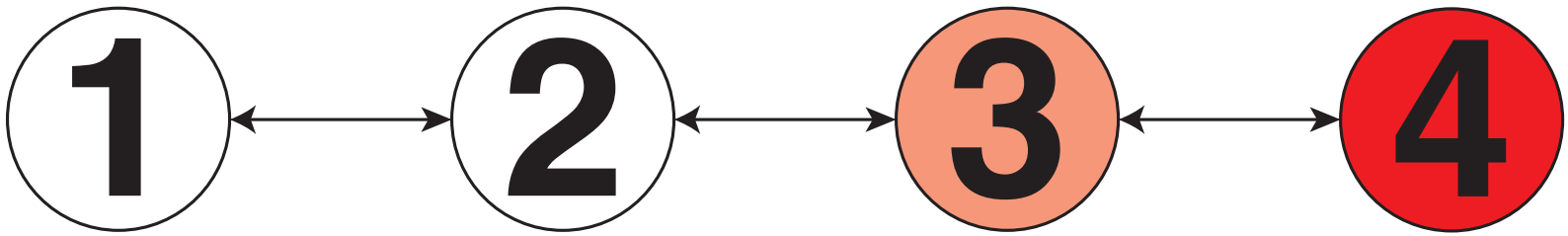
$$D_i = \begin{cases} \text{Di(rect)}: & Z_i = 1 \\ \text{In(direct)} & Z_{i\pm 1} = 1 \\ \text{Co(ntrol)} & Z_i = Z_{i\pm 1} = 0. \end{cases}$$

- There is nothing particularly special about this model, except for its parsimony. Arbitrarily complex exposure models are possible.
- Let's visualize this.


1
2
3
4


1
2
3
4


1
2
3
4


1
2
3
4

Summarizing:

		Unit #			
		1	2	3	4
Rand. #	1	1	0	0	0
	2	0	1	0	0
	3	0	0	1	0
	4	0	0	0	1
Design Z_i					

$\longrightarrow$

		Unit #			
		1	2	3	4
Rand. #	1	Di	In	Co	Co
	2	In	Di	In	Co
	3	Co	In	Di	In
	4	Co	Co	In	Di
Exposure D_i					

We can figure out the exact probabilities that each of the four units would be in each of the exposure conditions:

		Unit #			
		1	2	3	4
Rand. #	1	Di	In	Co	Co
	2	In	Di	In	Co
	3	Co	In	Di	In
	4	Co	Co	In	Di
Exposure D_i					

	Unit #			
	1	2	3	4
Direct	0.25	0.25	0.25	0.25
Indirect	0.25	0.50	0.50	0.25
Control	0.50	0.25	0.25	0.50
Probabilties $\pi_i(D_i)$				

Neyman-Rubin model: *potential outcome* associated with each exposure, but "fundamental problem of causal inference" in that we observe only one potential outcome per unit.

- If unit i receives exposure d_k, outcome is $Y_i(d_k)$.

	Unit #				
	1	2	3	4	Mean
Direct	5	10	10	3	7
Indirect	0	3	3	2	2
Control	1	3	6	2	3
Potential outcomes $Y_i(D_i)$					

- Average causal effect: $\tau(d_k, d_l) = \frac{1}{N} \sum_{i=1}^{N} [Y_i(d_k) - Y_i(d_l)]$.
- E.g., $\tau(Direct, Control) = \frac{1}{N} \sum_{i=1}^{N} [Y_i(Direct) - Y_i(Control)] = 4$.

- Unequal probability design provides a natural, and design-unbiased estimator. The Horvitz-Thompson (HT) estimator:

$$\hat{\tau}_{HT}(d_k, d_l) = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{\mathbf{I}(D_i = d_k)}{\pi_i(d_k)} Y_i(d_k) - \frac{\mathbf{I}(D_i = d_l)}{\pi_i(d_l)} Y_i(d_l) \right]$$

- Unbiasedness is very easy to see.

$$\mathrm{E}\left[\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\mathbf{I}(D_i = d_k)}{\pi_i(d_k)}Y_i(d_k) - \frac{\mathbf{I}(D_i = d_l)}{\pi_i(d_l)}Y_i(d_l)\right]\right] =$$

$$\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\mathrm{E}\left[\mathbf{I}(D_i = d_k)\right]}{\pi_i(d_k)}Y_i(d_k) - \frac{\mathrm{E}\left[\mathbf{I}(D_i = d_l)\right]}{\pi_i(d_l)}Y_i(d_l)\right] =$$

$$\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\pi_i(d_k)}{\pi_i(d_k)}Y_i(d_k) - \frac{\pi_i(d_k)}{\pi_i(d_l)}Y_i(d_l)\right] =$$

$$\frac{1}{N}\sum_{i=1}^{N}[Y_i(d_k) - Y_i(d_l)] = \tau(d_k, d_l)$$

- Unbiasedness follows from very clear assumptions:
- How was the randomization administered? (known)
- What is the exposure model? (assigned by analyst)
- These assumptions are always being made, although often obscured and/or inconsistent with the experimental design
- Here, design and assumptions directly motivate the estimator

- E.g., for the first randomization $\mathbf{z} = (1, 0, 0, 0)$, we would observe:

Y_i	5	3	6	2
Z_i	1	0	0	0
D_i	Di	In	Co	Co
$\pi_i(D_i)$	0.25	0.50	0.25	0.50

- HT estimator:

$$\hat{\tau}_{HT}(Di, Co) = \frac{1}{4}\left[\frac{5}{0.25} - \left(\frac{6}{0.25} + \frac{2}{0.50}\right)\right] = -2$$

.

- Can also look at the difference in means estimator (logically equivalent to an OLS regression of the outcome on treatment dummies):

$$\hat{\tau}_{DM}(Di, Co) = \frac{5}{1} - \frac{6+2}{2} = 1$$

.

- So let's see how the HT estimator performs against the difference in means estimator

Across all randomizations,

		Diff. in Means		$\widehat{\tau_{HT}}(d_k, d_l)$	
Rand. #	1	1.00	-1.00	-2.00	-5.50
	2	8.00	-0.50	9.00	0.50
	3	9.00	1.50	9.50	3.00
	4	1.00	1.00	-0.50	-2.00
	E[.]	4.75	0.25	4.00	-1.00
	Bias	0.75	1.25	0.00	0.00
		$\tau(Di, Co)$	$\tau(In, Co)$	$\tau(Di, Co)$	$\tau(In, Co)$

- The difference in means / OLS estimator is badly biased – in fact, in, expectation, it even gets the sign wrong for the indirect effect
- Not just a small sample problem – bias even in asymptopia.

Inference:

$$\mathrm{Var}\left(\widehat{\tau_{HT}}(d_k, d_l)\right) = \frac{1}{N^2}\Big\{\mathrm{Var}\left[\widehat{Y_{HT}^T}(d_k)\right] + \mathrm{Var}\left[\widehat{Y_{HT}^T}(d_l)\right] - 2\mathrm{Cov}\left[\widehat{Y_{HT}^T}(d_k), \widehat{Y_{HT}^T}(d_l)\right]\Big\},$$

where,

$$\mathrm{Var}\left[\widehat{Y_{HT}^T}(d_k)\right] = \sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Cov}\left[\mathbf{I}(D_i = d_k), \mathbf{I}(D_j = d_k)\right]\frac{Y_i(d_k)}{\pi_i(d_k)}\frac{Y_j(d_k)}{\pi_j(d_k)}$$

$$\mathrm{Cov}\left[\widehat{Y_{HT}^T}(d_k), \widehat{Y_{HT}^T}(d_l)\right] = \sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Cov}\left[\mathbf{I}(D_i = d_k), \mathbf{I}(D_j = d_l)\right]\frac{Y_i(d_k)}{\pi_i(d_k)}\frac{Y_j(d_l)}{\pi_j(d_l)}$$

- Young's inequality provides approximations for unidentified components, and estimation proceeds using Horvitz-Thompson style estimator.
- In expectation, these approximations are conservative; and unbiased under sharp null hypothesis of no effect (for many designs).
- Asymptotic normality / conservative confidence intervals follow from restrictions on clustering.
- The paper contains "model-assisted" refinements for covariance adjustment, weight stabilization and constant effects variance estimation.

Example: Paluck and Shepherd (2012)

- (Rough) design:
  - Measured connections between 291 students with predeployment survey (via listing of friends)
  - Identified 83 "key" individuals, randomized 30 into attending an anti-bullying program
  - Measured behavioral and attitudinal outcomes for all 291 students
- How to analyze?
  - Interested in both *direct* (effects of attending program) and *indirect* effects (effects of peers attending program)
  - Heterogeneous (and sometimes zero) probabilities of exposure, implicit clustering
  - Outcome variable (for illustration): teacher evaluations of behavior (higher score = worse behavior)

Example: Paluck and Shepherd (2012)

- Consider the following exposure model:
  - Control: Not attending program, no peers in program
  - Direct: Attending program, no peers in program
  - Indirect: Not attending program, peers in program
  - Combined: Attending program, peers in program
- Some complexities. Effects estimated will be "local" average treatment effects.
- Can use more/less complex exposure models

Example: Paluck and Shepherd (2012)

Exposure	Naive (Diff-in-Means)	Regression (Fixed Effects)	HT (Ours!)
Direct	-0.775	-0.752	-1.400
(SE)	(0.793)	(0.927)	(1.133)
Indirect	-0.382	-0.648	-0.607
(SE)	(0.434)	(0.596)	(1.106)
Combined	-1.331	-1.663	-1.792
(SE)	(0.956)	(1.220)	(1.617)

Anticipating some concerns, sensitivity analysis, & implications.

Concern: "But you're still specifying an exposure model! What if you don't believe it?"

- We *always* have to specify an exposure model if we want to define causal effects.
- But! The framework permits exposure models of arbitrary generality.
- By definition, there exists a finite (but potentially very large) set of distinguishable exposure models that may be associated with any randomization scheme.
- These models can be nested in any arbitrary order. We can permit an arbitrarily large number of forms of interference in a series of nested models, all the way down to allowing exposure to be defined by the entire vector $\mathbf{Z}$.
- We can even reject null hypotheses of no (or fewer forms of) interference if we pick up on effects.

Sensitivity analysis?

- Sensitivity analysis really isn't at play here, since causal parameters are not well defined if the exposure model is incorrect (or, rather, incomplete).
- Without theory, we don't have an estimand.
- But many, many theories may be jointly implemented in a complex exposure model. Even if some exposures are irrelevant, it's only an issue of efficiency.
- "Sensitivity analysis" is then permitting additional levels/types of exposure.

Some other thoughts / extensions

- Principal strata?
  - No reason why we couldn't *estimate* traits of the exposure model, even based on information revealed by treatment assignment.
- Incomplete network data?
  - Imputation model, integrating over θ
- Observational studies?
  - If we can *estimate* the treatment assignment mechanism, then simple enough to specify an exposure model again.
- SUTVA?
  - Under proper specification, exposure model implies no interference.
  - Consistency assumption still necessary for external validity. With consistency, we satisfy SUTVA.

Conclusion:

- Exogeneity does not imply unbiasedness.
- Equal probability of assignment does not imply equal probability of exposure.
- Simple, nonparametric assumptions can clarify both questions and answers.